ANALYSIS OF PHYTOCHROME SIGNALING PATHWAYS IN *Zea mays*

BY

INDRAJIT KUMAR

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Biology
with a concentration in Physiological and Molecular Plant Biology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2012

Urbana, Illinois

Doctoral Committee:

Associate Professor Matthew Hudson, Chair
Associate Professor Stephen Moose
Assistant Professor Steven Clough
Associate Professor Ray Ming

Abstract

Phytochromes are photoreceptors that mediate red and far-red light signaling and control growth and development of plants in response to different light conditions. The genome of Arabidopsis, a dicot, contains five members of the phytochrome family, phyA-phyE. In contrast, genomes of monocots sequenced so far contain only three members, phyA-phyC, although polyploids may have multiple homeologs of each. Studies in Arabidopsis have shown that phytochromes primarily transduce light signals via interaction with PIFs (phytochrome interacting factors), a sub-family of bHLH proteins. The recent whole genome duplication in maize may have led to divergence in the structure and function of phytochrome mediated signaling pathways. Here, I found that ZmphyB1, one of the phyB homeologs, interacts with a bHLH protein similar to AtPIF3. Maize PIF3 exists as two homeologs, ZmPIF3e and ZmPIF3-P, and both homeologs were found to interact with the C-terminal domains of maize phyA2, phyB1 and phyB2. ZmPIF3e, however, interacts with the full length phyB1 but not the full length phyB2, suggesting a potential new function of maize phyB2. A cross-species interaction between phytochromes and PIFs of maize and Arabidopsis was also observed, suggesting a functional conservation of phytochrome and PIF structural contacts in the two evolutionarily distant species. In addition to the PIF family, several bHLH proteins play significant roles in plant growth and development. A genome-wide bioinformatic analysis revealed that the maize genome contains at least 197 putative bHLH encoding genes. Phylogenetic analysis and comparison of bHLH domain sequences, intron patterns inside bHLH domain, presence of non-bHLH motifs, and potential DNA binding features of maize bHLH proteins indicate functional conservation of bHLH proteins across maize, rice and Arabidopsis. However, some of the bHLH subfamilies contain non-bHLH motifs typical of other subfamilies and thus indicate a domain/motif shuffling during evolution leading to potential acquired new functions. More than 50% of maize bHLH genes exist as homeologs and again may provide a source of recently acquired new functions or subfunctionalization of duplicated bHLH genes.

Acknowledgements

My PhD would always be a treasured journey to me. To begin with, I would like to express my deepest gratitude to my advisor Dr Matthew Hudson for his unceasing encouragement and faith in me, without which this journey would not have been possible. Matt has been a great mentor and inspiration to me. I would also take this opportunity to thank Dr Karen Hudson for her expert suggestions and help in troubleshooting throughout my PhD.

I would also like to thank Dr Steve Moose and Dr Steve Clough for their key suggestions and also being gracious to let me use their lab facilities. I am deeply grateful to Dr Kankshita Swaminathan, Dr Gopal Battu and Dr Tong Geon Lee for their continuous help and encouragement. My sincere thanks also goes to all of my current and former lab members for their assistance and support.

I would like to thank my parents and family members for their love and care. Finally, this work would not have been possible without my wife Devi, who has served as an anchor in my life. I am grateful to the almighty that she became a part of my life.

Table of Contents

Phytochrome Nomenclature

PHY: Phytochrome gene (Arabidopsis)

Phy: Phytochrome gene (Maize)

PHY: Phytochrome apoprotein

phy: Phytochrome holoprotein (with conjugated chromophore)

Chapter 1

General Introduction

Photoreceptors in plants

Perception of light signal is essential for growth and development in plants. Light is not only the primary source of energy, but it has also been shown to regulate various developmental processes in plants e.g. seed germination, leaf expansion, phototropism, flowering time, grain filing, dormancy, circadian rhythms, and shade avoidance (Deng and Quail 1999; Jiao et al. 2007; Franklin and Quail 2010). Defense related pathways have also been shown to be strongly influenced by light signals (Genoud et al. 2002; Zeier et al. 2004). Plants measure the presence or absence, spectral quality (wavelength), fluence rate, direction and duration of the surrounding light conditions.

Light signals are perceived through four different families of photoreceptors, phytochromes (phys), cryptochromes (crys), phototropins (phots) and the ZTL family which includes ZTL (ZEITLUPE, LOV/F-box/Kelch-repeat proteins), FKF (FLAVIN-BINDING KELCH REPEAT F-BOX) and LKP2 (LOV KELCH-REPEAT PROTEIN 2) (Kim et al. 2007; Sawa et al. 2007; Moglich et al. 2010). Phytochromes are primarily involved in absorbing red (R) and far-red (FR) wavelengths. The cryptochrome, phototropin and ZEITLUPE family of photoreceptors mediate blue light/UV-A (320-500nm) signaling. Recently, a new photoreceptor UVR8 (UV RESISTANCE LOCUS 8) has been reported which perceives UV-B (282-320nm) wavelengths (Rizzini et al. 2011). These photoreceptors perceive light signals and transduce intracellular pathways which in turn modulate gene expression resulting in adaptive changes at the cellular and organismic level.

Phytochromes

The Phytochrome family of photoreceptors has been extensively studied in *Arabidopsis thaliana* (Franklin and Quail 2010) and consists of five members: phyA, phyB, phyC, phyD and phyE (Sharrock and Quail 1989; Clack et al. 1994). These are biliproteins which typically absorb R and FR light. Phytochromes form a dimer with each monomer consisting of an apoprotein (~125KDa) covalently attached to a light absorbing linear tetrapyrrole chromophore, phytochromobilin (PϕB) (Rockwell et al. 2006). Phytochromes exist in two reversible conformations, each with distinct spectroscopic and functional properties: the red light absorbing Pr form which is biologically inactive and the FR light absorbing Pfr form (biologically active). Phytochromes are produced in the Pr form. Fig 1.1 shows the absorption spectra of a typical plant phytochrome. Upon absorption of R, the Pr form gets converted to biologically active Pfr and conversely the active Pfr form reverts back to Pr by absorbing FR or by slow dark reversion (Rockwell et al. 2006) (Fig 1.2). Pr to Pfr conformational change leads to exposure of a NLS (nuclear localization signal) located in the C-terminal domain and thus Pfr gets translocated in to the nucleus (Yamaguchi et al. 1999; Nagatani 2004; Chen et al. 2005) where it forms speckles or nuclear bodies (Kircher et al. 1999; Yamaguchi et al. 1999). It has also been shown that nuclear translocation of phyA and phyB is required for their biological functions (Huq et al. 2003; Matsushita et al. 2003; Hiltbrunner et al. 2006; Toledo-Ortiz et al. 2010). However, recently it has been demonstrated that Arabidopsis phyB (Pfr) present in the cytosol is also functional and regulates the translation of *PORA* (protochlorophyllide reductase) mRNA, a key chlorophyll biosynthetic gene, by interacting with a cytosolic RNA binding C3H-type Zn finger protein PENTA1 (PNT1)(Paik et al. 2012). phyA has also been shown to possess cytosolic

functions like enhanced phototropism in red light and controlling negative gravitropism in blue light (Rosler et al. 2007).

Inside the nucleus, phytochromes interact with transcription factors, regulating light responsive genes. About 10% of the Arabidopsis transcriptome has been suggested to be under control of phytochromes (Quail 2002). One primary mechanism of phytochrome signaling has been found to be physical interaction of Phys (Pfr form) with a subfamily of bHLH transcription factors, PIFs (PHYTOCHROME INTERACTING FACTORS) which are negative regulators of light signaling (Ni et al. 1998; Ni et al. 1999; Monte et al. 2007; Shin et al. 2007; Leivar et al. 2008). This interaction leads to phosphorylation and subsequently proteasome mediated degradation of PIFs which in turn results in a change in gene expression profile activating photomorphogenesis (Bauer et al. 2004; Al-Sady et al. 2006; Leivar et al. 2008; Lorrain et al. 2008; Shen et al. 2008; Leivar et al. 2009; Shin et al. 2009).

Phytochrome Domain Structure

Plant phytochromes consist of an N-terminal chromophoric (~70 KDa) and a C-terminal dimerization moiety (~55 KDa) which can be further divided into subdomains (Fig 1.3a). The N-terminal domain comprises four subdomains: the N-terminal extension, PAS (Per/Arnt/Sim) domain, the GAF (cGMP phosphodiesterase/adenyl cyclase/FhlA) and a phytochrome domain (Montgomery and Lagarias 2002; Nagatani 2010). A cysteine residue inside the GAF domain covalently binds to the chromophore (PϕB) and together with the PAS and PHY domain, constitutes the core module responsible for photoreception (Essen et al. 2008; Nagatani 2010).

Crystal structure of a bacterial phytochrome revealed a light sensing knot like structure between PAS and GAF domains and a tongue like structure connecting the Phy domain to the knot (Fig 1.3b) (Nagatani 2010). This photosensory core has been found to be conserved in phytochromes from broad range of organisms including green algae, diatoms, cyanobacteria and fungi (Rockwell and Lagarias 2010) (Fig 1.3a). The C-terminal half of phytochrome contains two PAS domains followed by a histidine-kinase-related-domain (HKRD) (instead of an authentic HK domain present in bacterial phytochromes) and has been shown to be involved in nuclear localization and dimerization (Quail 1997). Analyses of various mutated versions of the phyB C-terminal domain indicated that the PRD (PAS related domain) in Arabidopsis phyB is essential and also sufficient for nuclear localization indicating the presence of potential NLS (nuclear localization signal) in the C-terminal region (Chen et al. 2005). The N-terminal domain of Arabidopsis phyB interacts with the C-terminal PRD region. This interaction is much stronger in the Pr form compared to the Pfr form and it has been suggested that a structural rearrangement in Pfr leads to exposure of the NLS in the C-terminal region, which otherwise is not exposed in the Pr form (Chen et al. 2005). Based upon structural studies, it was proposed that a Z to E isomerization across a methine bridge (C15=C16) takes place leading to photoconversion from Pr to Pfr (Fig 1.4) (Rockwell et al. 2006). However, recent NMR studies on another bacterial Cph1 phytochrome (Pfr) indicated a 90˚ rotation of ring around C4=C5 double bond instead as the mechanism of photoconversion (Fig 1.4) (Ulijasz et al. 2010). The HKRD domain is thought to be an evolutionary remnant as it lacks a critical histidine residue present in active histidine-kinase domain in bacterial phytochromes (Boylan and Quail 1996). Earlier, it was shown that the C-terminal domain is essential for downstream signaling as some

point mutations in the C-terminal domain of both phyA and phyB in Arabidopsis led to loss of their biological activity (Quail et al. 1995; Quail 1997). However, a dimer of the N-terminal domain of Arabidopsis phyB was later shown to perceive the light signal and transduce to downstream components at much higher sensitivity than the full length phyB (Matsushita et al. 2003; Oka et al. 2004; Palagyi et al. 2010). It was then suggested that the N-terminal domain of phyB transduces the light signal and the C-terminal domain attenuates phyB activity (Matsushita et al. 2003).

Phytochrome signaling and bHLH transcription factors in Arabidopsis

The five Arabidopsis phytochromes, phyA-phyE, have been classified into two groups based upon their stability in light: type I (light labile) which includes phyA and type II (light stable) which includes phyB-phyE (Sharrock and Quail 1989). phyA has been found to be most abundant in dark grown seedlings. When exposed to R or white light (W), the phyA protein level drops rapidly. In contrast, phyB is quite abundant in light grown plants. The phyC, phyD and phyE proteins are neither abundant in etiolated plants, nor strongly attenuated in response to light (Clack et al. 1994; Hirschfeld et al. 1998; Sharrock and Clack 2002). Based upon protein sequence analysis, these phytochromes are clustered into three subfamilies: phyA/phyC, phyB/phyD and phyE (Clack et al. 1994). Orthologs of Arabidopsis *PHY* genes are present in most of the higher plants (Clack et al. 1994; Mathews and Sharrock 1997).

Phytochromes transduce the light signal to the transcriptional network via a direct physical interaction with a small group of basic helix-loop-helix (bHLH) transcription factors, also termed

as Phytochrome Interacting Factors (PIFs) (Duek and Fankhauser 2005; Castillon et al. 2007). The bHLH family is defined by the signature bHLH domain consisting of 60-70 amino acids and the domain is arranged in typical bifunctional regions: a basic region of 13-17 residues located at the N-terminal end of the domain and a HLH (helix-loop-helix) region at the C-terminal end (Fig 1.5a). The basic region is enriched with basic residues and is involved in DNA binding. The HLH region is constituted by two conserved α- helices rich in hydrophobic residues that are separated by a variable length of loop (Ferre-D'Amare et al. 1993) which facilitate homo or hetero-dimer formation by protein-protein interaction (Massari and Murre 2000). Each monomer has been shown to bind to half of the DNA recognition sequence primarily via the basic region (Ma et al. 1994; Shimizu et al. 1997). bHLH proteins recognize a core DNA sequence motif which is a consensus six nucleotide sequence known as the E-box (5'-CANNTG-3'). One of the most common types of E-box consists of sequence 5'-CACGTG-3'and is known as G-box. Many plant bHLH proteins have been shown to have a preferential recognition and binding to G-box (Martinez-Garcia et al. 2000; Huq and Quail 2002; Huq et al. 2004). The basic region of the bHLH domain contains certain conserved residues which dictate the recognition of core consensus DNA sequence and some other residue provide specificity for the type of E-box e.g. G-box, N-box (Robinson et al. 2000).

Recently, there have been several studies characterizing the bHLH superfamily in plants including Arabidopsis and rice, and the PIFs belong to subfamily 15 of the bHLH superfamily (Heim et al. 2003; Toledo-Ortiz et al. 2003; Li et al. 2006; Pires and Dolan 2010). The Arabidopsis PIF group consists of PIF1 (PIL5), PIF3, PIF4, PIF5, PIF6 (PIL2), PIF7 and PIF8, all of which have been shown to interact with phyB (Fig 1.5b) (Khanna et al. 2004; Leivar and Quail

2011). Other members of this subfamily such as PIL1, HFR1 and SPT do not interact with phytochrome, however they play a role in phytochrome mediated photomorphogenesis (Khanna et al. 2004; Castillon et al. 2007). HFR1 contains an atypical basic domain lacking the two conserved residues essential to bind the G-box motif (Fairchild et al. 2000). PIF3 was the first member of the PIF subfamily to be identified as a phyB interacting protein through yeast two hybrid screening, and was later shown to bind to the Pfr form of phyA and phyB (Ni et al. 1998; Ni et al. 1999). PIF3 binds to phyB with greater affinity than phyA in Arabidopsis and the PAS domain of PIF3 was shown to be the major contributor for this interaction (Zhu et al. 2000). The N-terminus of phyB contains a 37 amino acid segment which is not present in phyA. A deletion mapping study indicated that the 37 aa segment of phyB contributes to the higher binding affinity of phyB to PIF3 (Zhu et al. 2000). Besides the typical bHLH, the PIFs also contain a conserved Active Phytochrome B binding (APB) motif in the N-terminal region, shown to be essential and sufficient for interaction with light activated phyB (Khanna et al. 2004). All of the PIFs were shown to interact with phyB (Pfr) (Huq and Quail 2002; Huq et al. 2004; Oh et al. 2004; Leivar et al. 2008). PIF1, PIF3 and PIF6 possess higher affinity to phyB (Pfr) than PIF4 and PIF5 (Huq and Quail 2002; Huq et al. 2004; Khanna et al. 2004). PIF1 and PIF3 also bind to phyA although PIF1-phyA interaction is much stronger (Ni et al. 1998; Huq et al. 2004). Not surprisingly, PIF1 and PIF3 contain an Active Phytochrome A binding (APA) motif although it is not well conserved (Al-Sady et al. 2006; Shen et al. 2008). As bHLH proteins are known to form dimers, PIFs have also been found to form homodimers and heterodimers. PIF3 can form a homodimer as well as a heterodimer with PIF4 (Toledo-Ortiz et al. 2003). PIF3-PIF4 heterodimers are also capable of binding to the G-box DNA elements (Toledo-Ortiz et al. 2003).

PIF3 also heterodimerizes with HFR1, an atypical bHLH protein which function in phyA and cryptochrome mediated far-red and blue light signaling pathways (Fairchild et al. 2000; Duek and Fankhauser 2003).

The activity of PIFs is regulated at different levels and mainly via the interaction with light – activated phytochromes which leads to phosphorylation and degradation of PIFs as shown for PIF1, PIF3, PIF4 and PIF5 (Park et al. 2004; Al-Sady et al. 2006; Shen et al. 2007; de Lucas et al. 2008; Lorrain et al. 2008; Shen et al. 2008). A rapid light induced ubiquitination of PIF1, PIF3 and PIF5 prior to degradation has been shown which indicates similar regulatory mechanism for all PIFs (Al-Sady et al. 2006; Shen et al. 2007; Shen et al. 2008). This mechanism of regulation leads to high level of PIFs in dark and a rapid reduction when exposed to light with high R:FR. PIF1 and PIF3 also declines significantly under FR light however high levels of PIF4 and PIF5 are retained under FR light (Bauer et al. 2004; Lorrain et al. 2008; Shen et al. 2008). Interestingly, PIF7 levels are not light regulated even though it interacts with phyB (Leivar et al. 2008). Recently, PIF7 has been shown to be involved in shade avoidance syndrome (SAS) response where PIF7 accumulates in its dephosphorylated state under shade conditions, binds to auxin biosynthetic genes and thus increases auxin synthesis leading to shade induced growth (Li et al. 2012).

For some PIFs, heterodimerization with other transcription factors is another mode of activity regulation where the hetero-dimers lack the DNA binding ability (de Lucas et al. 2008; Feng et al. 2008). PIF3 and PIF4 heterodimerize with DELLA family members resulting in a crosstalk between hormone and light mediated growth (de Lucas et al. 2008; Feng et al. 2008). PIF4 and

PIF5 activity has also been shown to be inhibited by heterodimerization with HFR1, an atypical bHLH protein which is present in high level under shade (low R:FR) conditions (Sessa et al. 2005; Hornitschek et al. 2009).

PIFs have been shown to be involved in variety of functions. PIF1 represses seed germination in darkness by inhibiting key gibberellin (GA) biosynthetic genes *GA3ox1* and *GA3ox2* and promotes GA catabolic gene *GA2ox2* expression thus minimizing the GA level (Oh et al. 2007). At the same time, PIF1 also promotes expression of abscisic acid (ABA) biosynthetic genes *ABA1*, *NCED6* and *NCED9* and inhibits expression of the ABA catabolic gene *CYP707A2,* thereby leading to high levels of ABA (Oh et al. 2007). PIFs (PIF7, PIF3, and PIF4) also possess a major role in seedling de-etiolation under prolonged red light by regulating the phyB levels via a negative feedback loop mechanism (Al-Sady et al. 2008; Leivar et al. 2008). Recently, PIFs have been shown to activate COP1 (CONSTITUTIVE PHOTOMORPHOGENIC 1) mediated ubiquitination and subsequent degradation of phyB (Jang et al. 2010).

Shade avoidance syndrome

Several plant species including agricultural crops are often found to be in situations with limited resources leading to competition between individuals. Light being a critical resource, plants elicit various responses to adapt to and compete for the ambient light environment. Under decreased light quantity and quality (low R:FR and low blue light) such as vegetation shade or high plant density, plants elicit morphological and physiological responses known as the Shade Avoidance Syndrome (SAS). Fig 1.7 shows the spectral composition of different light

environments. These responses include elongation of stem, increased apical dominance reduced branching, decreased leaf area (Smith and Whitelam 1997) which help plants come out of canopy and overcome competitors. Under persistent shade condition, plants accelerate flowering and early seed production (Halliday et al. 1994). Physiological responses include a redistribution of auxin, increased ethylene production and an acceleration of flowering (Smith 1995; Morelli and Ruberti 2000; Vandenbussche et al. 2003). Light transmitted through any plant canopy contains attenuated red and blue wavelengths due to absorption by chlorophyll and carotenoids and utilization for photosynthesis. Thus, FR and green wavelengths get enriched leading to reduced R:FR ratio. Not surprisingly, SAS involves interaction between phytochrome and UV/blue light photoreceptor signaling (Franklin 2008). Recently, it was shown that the abundant green light under canopy shade also informs plants of shade condition and induces shade avoidance responses (Zhang et al. 2011). Under high R:FR conditions, shade avoidance responses are primarily suppressed by phyB mediated signaling (Nagatani et al. 1991; Somers et al. 1991). Various phytochrome mutant analyses revealed redundant roles of phyB, phyD and phyE (Aukerman et al. 1997; Devlin et al. 1998; Devlin et al. 1999; Franklin et al. 2003). phyB and phyE have been shown to repress expression of a homeodomain leucine zipper (HD Zip) transcription factor *ATHB2*, a known promoter of SAS responses (Carabelli et al. 1996; Franklin et al. 2003). *PIL1* and *HFR1,* both bHLH transcription factor coding genes, are also early acting genes during SAS (Salter et al. 2003; Sessa et al. 2005).

Phytochromes in monocots

So far, phytochrome signaling has been studied extensively in dicots especially Arabidopsis however, little is known about it in monocots. Maize, rice, wheat and sorghum belong to agronomically and economically important monocot plants which are staple food crops throughout the world. There have been few studies related to phytochrome mediated signaling in rice (Takano et al. 2001; Takano et al. 2009; Osugi et al. 2011; Todaka et al. 2012) and maize (Christensen and Quail 1989; Sheehan et al. 2004; Dubois et al. 2010) . Phytochrome has been shown to be involved in controlling agronomically important traits such as axillary stem (tiller) number in sorghum and rice (Foster et al. 1994; Kong et al. 2004; Kebrom et al. 2006) and stand density in maize (Maddonni et al. 2002; Fellner et al. 2003). Overexpression of phytochromes (phyA and phyB) has been repeatedly used to increase yield in densely-planted crops (Robson et al. 1996; Garg et al. 2006; Rao et al. 2011). Overexpression of the Arabidopsis *PHYA* gene in rice leads to increase in yield; however it also showed unwanted pleiotropic effects such as dwarf plants with reduced internode length and diameter (Garg et al. 2006). Ectopic expression of Arabidopsis *PHYB* in potato increased yield in a densely planted crop however plants were shown to have higher stomatal conductance, transpiration rate and photosynthesis rates. (Thiele et al. 1999; Boccalandro et al. 2003; Schittenhelm et al. 2004). In a recent study, introduction and overexpression of Arabidopsis *PHYB* gene in cotton showed a 35% increase in yield with pleiotropic effects like semi-dwarfism, decrease in apical dominance and increased boll size (Rao et al. 2011). Thus, manipulation of phytochrome mediated pathways has immense potential in improving crop yield and reversed shade avoidance if negative phenotypic effects can be avoided.

A similar approach can be applied in maize for improvement in yield, however a thorough understanding of phytochrome mediated signaling is required in order to avoid unwanted effects. Unlike Arabidopsis and other dicots, monocots such as rice and maize contain only three members in the phytochrome gene family: *PhyA*, *PhyB* and *PhyC* (Mathews and Sharrock 1996; Mathews and Sharrock 1997). Phytrochromes are present in single copies in rice and sorghum, in contrast, maize contains two homeologs of each phytochrome namely *PhyA1* and *PhyA2; PhyB1* and *PhyB2;* and *PhyC1* and *PhyC2* (Christensen and Quail 1989; Childs et al. 1997; Basu et al. 2000). Gene mapping studies in maize revealed that the *PhyB1* and *PhyB2* loci is present on chromosome 1S (1.03) and 9L (9.05/9.06) respectively. *PhyC1* and *PhyC2* was found on 1L (1.12) and 5S (5.02/5.03) and *PhyA1*-*PhyA2* on 1L and 5S respectively (Sheehan et al. 2004). As *PhyA1*, *PhyB1* and *PhyC1* genes are located on chromosome 1, these are thought to be derived from one ancestral genome and *phyB2* on chromosome 9 and *PhyA2*/*PhyC2* on chromosome 5 were most likely derived from another ancestral genome after an ancient allopolyploidization event (Gaut and Doebley 1997; Gaut 2001; Sheehan et al. 2004). Monocots and dicots are believed to have split ~140-200 Mya and since then they have been evolving independently. Therefore, it has been suggested that the functions of monocot and dicot phytochromes may not be totally conserved (Sheehan et al. 2004). In fact, several studies have indicated unique roles of Arabidopsis and rice phytochromes (Takano et al. 2001; Franklin and Whitelam 2004; Takano et al. 2005). In rice, phyA is involved in de-etiolation (inhibition of coleoptile elongation) via the VLFR (very low fluence response) to FR light however Arabidopsis phyA possesses a similar function (inhibition of hypocotyls elongation) via the high irradiance response (HIR) (Takano et al. 2005). Also, *Lhcb* (light-harvesting complex associated with

photosystem II) gene expression in rice is controlled by phyA in R/FR reversible manner, in contrast to the irreversible mode found in Arabidopsis (Takano et al. 2005). In rice, phyA and phyB are involved in R-light mediated inhibition of coleoptiles elongation, leaf length and internode elongation (Takano et al. 2005). In Arabidopsis, phyB possesses the primary role in controlling hypocotyl elongation in R light and R/FR photoreversible seed germination (Franklin and Quail 2010). Thus, there seems to be fundamental differences between monocot and dicot phytochrome-mediated pathways (Sheehan et al. 2004).

Maize PHYB1 and PHYB2 homeologs possess 94% identity at the protein level and they have been shown to have many redundant roles in controlling traits such as ear node height, full plant height, stem-diameter and leaf-sheath minus internode length (Sheehan et al. 2004). However, a subfunctionalization of phyB activities has also been reported where either phyB1 or phyB2 show major involvement in controlling particular traits such as flowering time regulation by phyB2, and inhibition of mesocotyl elongation in red light by phyB1 (Sheehan et al. 2004). It is likely that the downstream components of phyB1 and phyB2 may not be completely similar, leading to functional differences. For the same reason, identification of potential interacting partners of maize phyB proteins would be quite useful.

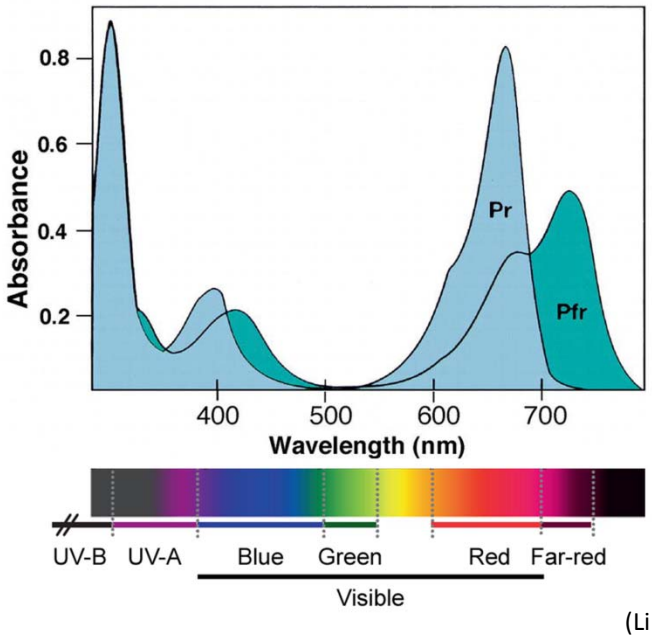
As discussed earlier, since the primary mode of mechanism of phytochromes has been shown to be interaction with PIFs (bHLH transcription factors), I was also interested in identifying the genes encoding the bHLH group of proteins and subsequently the PIFs across the maize genome by bioinformatic approaches. Arabidopsis phytochromes have been shown to possess redundant roles. In contrast, the maize genome contains only three phytochrome genes with

two homeologs each. This led to another question to determine the interaction possibilities of other maize phytochromes with PIF3 that is the primary interacting partner of phyB. Considering the fact that overexpression of Arabidopsis phyA/phyB in various transgenic crop plants led to potential increases in yield, it was also an important question to ask if maize phytochromes can interact with Arabidopsis PIF3 and *vice versa* despite the evolutionary distance. The information obtained from this study would be helpful in generating high yield varieties by modifying phytochrome mediated photoreception pathways.

As the shade avoidance syndrome is one of the limiting factors in improving crop yield by high density planting, information related to differential gene expression under shade conditions would also be significant. It has been shown that the EOD-FR (end-of-day FR) treatment results in similar physiological responses to those observed under canopy shade or crowding (Fankhauser and Casal 2004). Studies in Arabidopsis have shown that a pulse of far red light immediately preceding the dark period (EOD-FR treatment) results in a 10-fold increase in *ATHB2* expression after one hour of the treatment (Carabelli et al. 1996). Recently a physiological and genetic characterization of EOD-FR light response in maize was performed where *PhyB1* and *PhyB2* were shown to have redundant roles in mediating EOD-FR responses with *PhyB1* possessing the prominent role (Dubois et al. 2010). This study also revealed a major role of GA (Gibberellic Acid) in promoting mesocotyl elongation (Dubois et al. 2010). In order to identify differentially expressed transcripts that respond rapidly to EOD-FR treatment, I performed microarray analysis which would in turn help in identifying components of the shade avoidance response pathway in maize. Information obtained from these studies can prove to be

useful in yield improvement of maize and other crops by biotechnological manipulation in phytochrome pathway.

Figures and Tables



(Li et al. 2011)

Fig 1.1: Absorption spectra of plant phytochromes: The Pr form absorbs maximally at 660-670nm (R) and the Pfr form absorbs maximally at 725-735nm (FR).

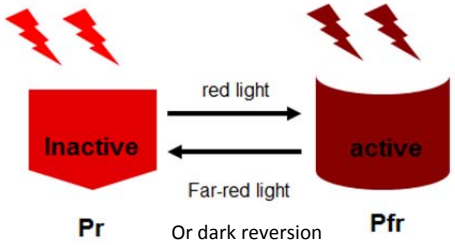
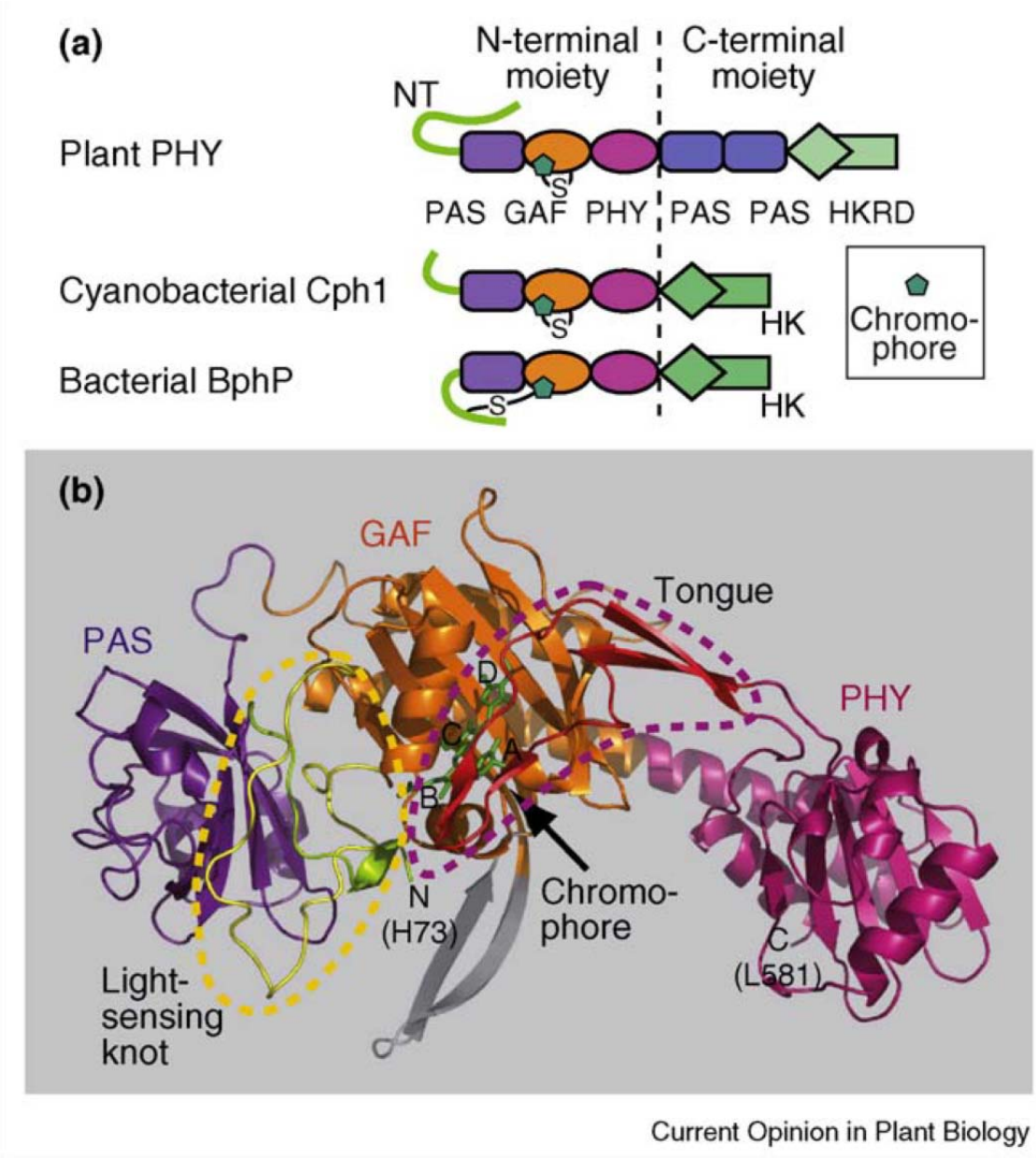


Fig 1.2: The photoconvertible Pr and Pfr forms of phytochrome

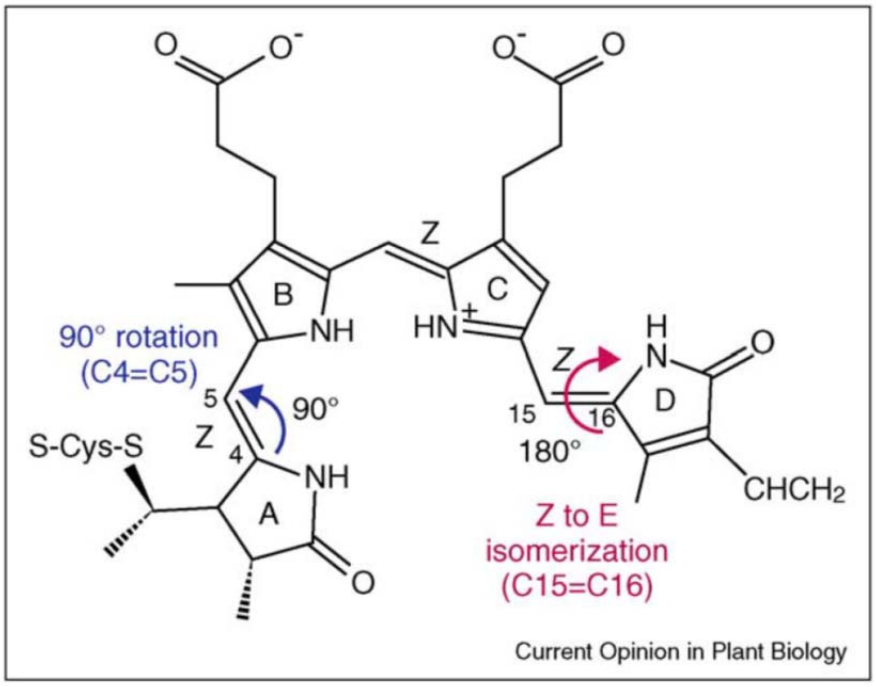


(Nagatani 2010)

Fig 1.3: Domain structures in Phytochromes.

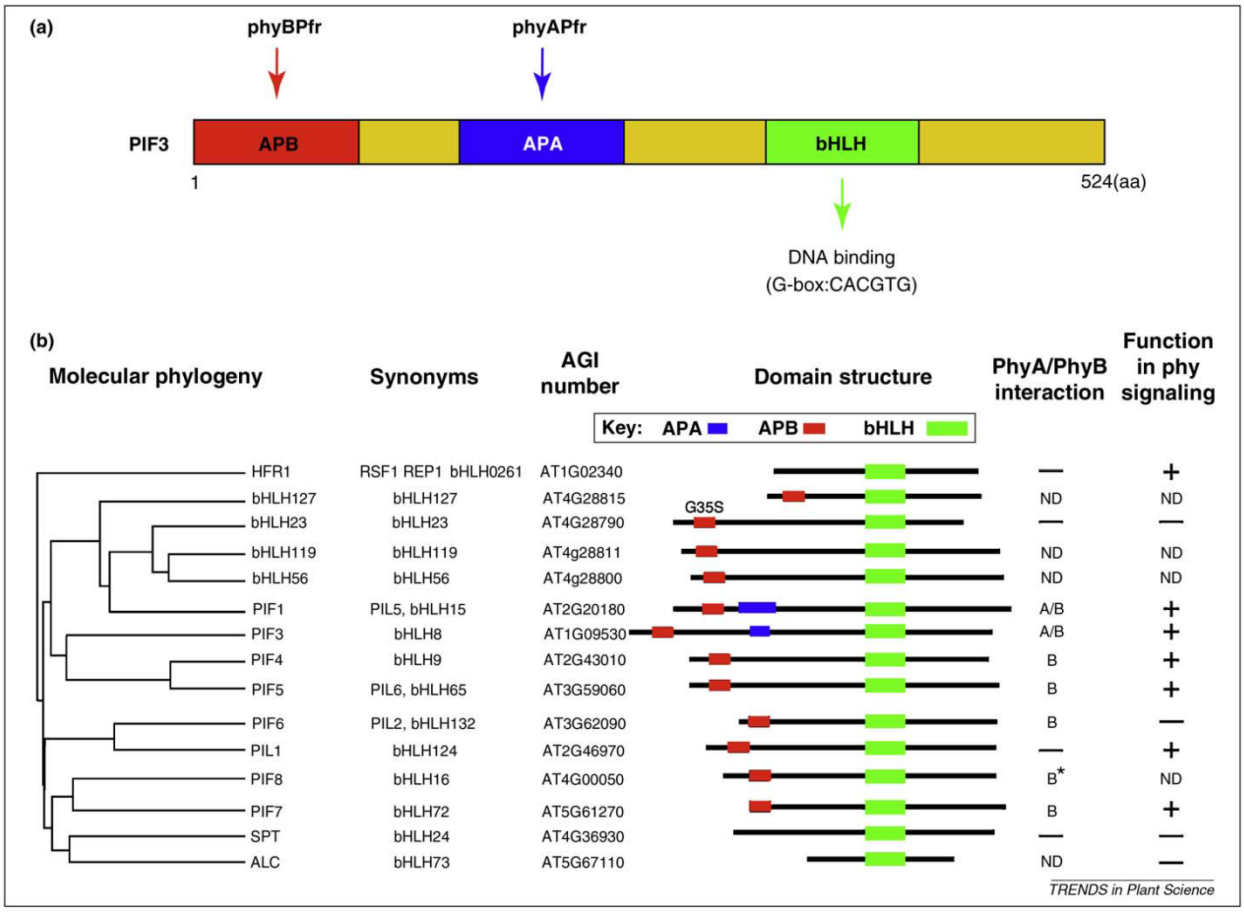
(a) Schematic representation of phytochrome domains in plant and bacterial phytochromes [NT: N-terminal extension; PAS: Per/Arnt/Sim domain; GAF: cGMP phosphodiesterase/adenyl cyclase/FhlA domain; HK: histidine kinase domain; HKRD: histidine kinase related domain]

(b) A 3-D homology model of Arabidopsis phyA N-terminal moiety showing the light sensing knot (yellow) and the tongue (red).



(Nagatani 2010)

Fig 1.4: Structure of phytochromobilin (phytochrome chromophore) showing the two proposed mechanisms for phytochrome photoconversion. The old hypothesis of Z to E isomerization around C15=C16 (red) and recently proposed C4=C5 90˚ rotation (blue).



(Leivar and Quail 2011)

Fig 1.5: The PIF subfamily of bHLH transcription factors in Arabidopsis. (a) Schematic representation of PIF3 domain structure showing the consensus bHLH domain, Active phyB Binding (APB) motif and Active PhyA binding motif (APA). (b) Subfamily 15 of Arabidopsis bHLH superfamily showing the molecular phylogeny, domain structure, phyA/phyB interaction status and functional status in phy signaling. [A/B represents interaction with Pfr form of phyA or phyB; (-) lack of interaction; ND: not determined; +/-/ND in the column for function represents positive evidence/lack of evidence or not determined]

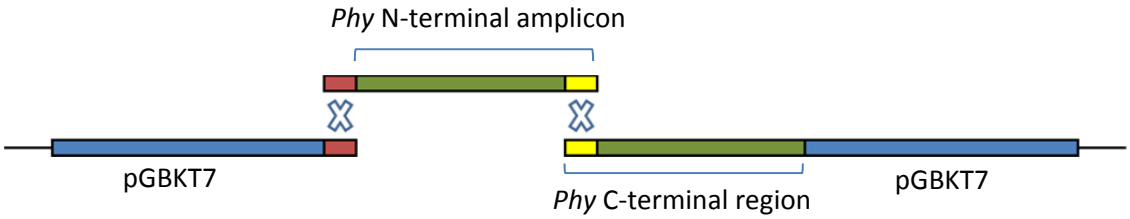
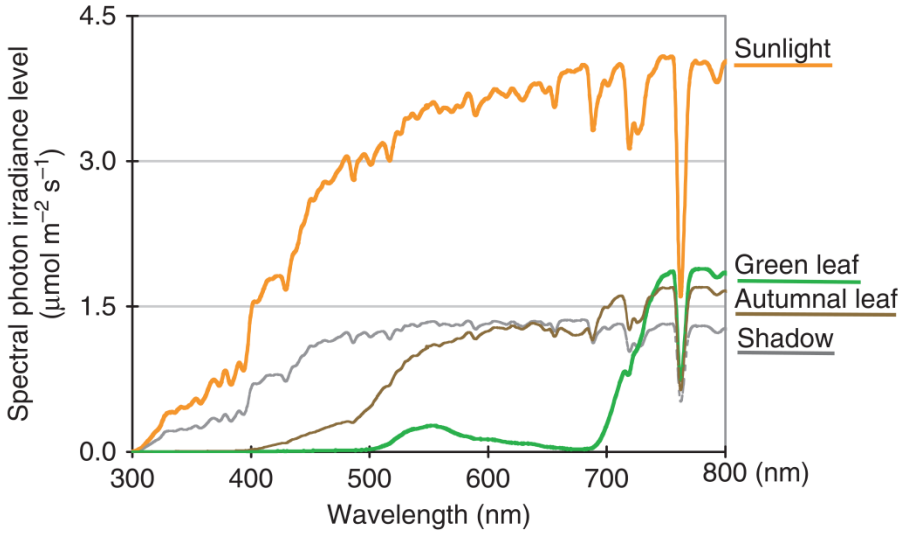


Fig 1.6: Schematic diagram of transformation assisted recombination method of cloning. The linearized plasmid vector (pGBKT7) containing the C-terminal region of phytochrome gene and the PCR amplicon of the N-terminal region of phytochrome gene were co-transformed into yeast. The overlapping regions (marked with red and yellow blocks) were introduced during PCR of phytochrome N-terminal region. Recombination events at the marked sites lead to fusion of N-terminal and C-terminal regions of the phytochrome gene.



(Kami et al. 2010)

Fig 1.7: A typical spectral irradiance of natural light environments- unfiltered sunlight, sunlight transmitted through green leaf and light brown autumn leaf and shadow of a building. Light transmitted through green leaf show very low R:FR.

Chapter 2

Phytochrome Interacting Partners in Maize

Abstract

Phytochromes transduce red and far red light signaling and affect several crop traits. The extent of functional conservation for phytochrome-mediated light signaling between dicots and monocots is not known but expected to show divergence due to past genome duplication events. Using the yeast two hybrid screening, an AtPIF3 homolog in maize was identified which showed interaction with maize PhyB1. The maize genome also possesses two PIF3 homeologs and I demonstrated that both PIF3 homeologs interact with maize phytochromes. My studies revealed cross-species interaction between Phy and PIFs of maize and Arabidopsis suggesting a functional conservation of the phytochrome systems in the two species. Similar to other plant systems, maize phytochrome also bound to PIF3 only in the active Pfr form but not the Pr form.

Introduction

Phytochrome family of photoreceptors are involved in sensing R (red) and FR (far-red) light and transducing intracellular signaling pathways modulating global gene expression in response to changing light environments. The Arabidopsis genome contains five members in the phytochrome family namely phyA-phyE (Sharrock and Quail 1989; Clack et al. 1994). Phytochromes exist as dimers with each monomer consisting of an apoprotein of ~125 KDa and a covalently attached linear tetrapyrrole chromophore, phytochromobilin (PϕB) that is responsible for absorbing light (Rockwell et al. 2006). Phytochromes have been shown to exist in two photoconvertible forms: the inactive Pr form and the biologically active Pfr form. Upon absorption of R light, the Pr form gets converted to biologically active Pfr and conversely the active Pfr form reverts back to Pr by absorbing FR light or by slow dark reversion (Rockwell et al. 2006). Pr to Pfr conformational change leads to exposure of a NLS (nuclear localization signal) located in the C-terminal domain of phytochromes and thus the Pfr form gets translocated into the nucleus (Yamaguchi et al. 1999; Nagatani 2004; Chen et al. 2005).

Inside the nucleus, Phytochromes have been shown to directly interact with several proteins including transcription factors. A yeast-two-hybrid screening revealed that the C-terminal domain of AtphyB interacts with a bHLH transcription factor PIF3 (Phytochrome Interacting Factor 3) (Ni et al. 1998). Subsequently, AtPIF3 was shown to co-immunoprecipitate with the Pfr form of phyB holophytochrome *in vitro* when reconstituted with the bacterial chromophore phycocyanobilin (PCB) (Ni et al. 1999). Subsequently, other PIFs including PIF1, PIF4, PIF5, PIF6,

PIF7 and PIF8 were also found to interact with phyB in a light dependent manner (Huq and Quail 2002; Huq et al. 2004; Oh et al. 2004; Leivar et al. 2008; Leivar and Quail 2011). All Arabidopsis PIF family members have been shown to possess a conserved motif near the N-terminus known as the Active Phytochrome B binding (APB) motif which provides the binding specificity of PIFs for phyB (Pfr) (Huq and Quail 2002; Huq et al. 2004; Oh et al. 2004; Leivar et al. 2008). APB motifs were also found to be essential and sufficient for PIF interaction with light activated phyB (Khanna et al. 2004). PIF1 and PIF3 also bind to phyA, although the PIF1-phyA interaction is stronger than PIF1-phyB (Ni et al. 1998; Huq et al. 2004). AtPIF1 and AtPIF3 contain an additional motif known as APA (Active Phytochrome A binding) motif responsible for phyA binding, however APA motifs are not well conserved (Al-Sady et al. 2006; Shen et al. 2008). The bHLH family of proteins in plants has been shown to have a preferential recognition and binding to G-box promoter motifs (5'-CACGTG-3') (Martinez-Garcia et al. 2000; Huq and Quail 2002; Huq et al. 2004). Indeed, AtPIF3 was found to bind G-box motifs present in the promoters of several light regulated genes and regulate their expression (Martinez-Garcia et al. 2000).

AtphyB has also been shown to interact with ARR4 (Arabidopsis Response Regulator 4) and this interaction stabilizes the active Pfr form thereby elevating the active phyB level (Sweere et al. 2001). The C-terminal domain of both AtphyA and AtphyB was shown to interact with PKS1 (phytochrome kinase substrate 1) in the cytoplasm (Fankhauser et al. 1999). This study also showed that both Pr and Pfr forms of both phytochromes bind to PKS1. PKS1 has been suggested to be a negative regulator of phytochrome activity and an *in vitro* study also showed that oat phyA can phosphorylate PKS1 (Fankhauser et al. 1999). Interestingly, PKS1 also

interacts with PHOT1 (Phototropin 1) and NPH3 (Non Phototropic Hypocotyl 3), both of which are involved in blue light mediated phototropism (Lariguet et al. 2006) indicating a complex cross talk between various photoreceptor pathways.

COP1 (Constitutive Photomorphogenic 1) is another phyB interacting protein which has been shown to bind to the N-terminal region of phyB leading to polyubiquitination and subsequent degradation of phyB (Jang et al. 2010). COP1 encodes for E3 ubiquitin ligase and has also been shown to physically interact and subsequently promote degradation of photomorphogenesis promoting transcription factors including HY5, HYH, LAF1 and HFR1 (Osterlund et al. 2000; Holm et al. 2001; Seo et al. 2003; Jang et al. 2005; Yang et al. 2005). It has been proposed that phytochromes may also be responsible for rapid initial inactivation of COP1 by a protein-protein interaction (Wang and Deng 2003). A simplified model of the phytochrome signaling pathway is shown in Fig 2.1 (Li et al. 2011). NDPK2 (Nucleoside-diphosphate kinase) has been shown to be another protein interacting with active Pfr form of phytochrome A and to act as a positive signaling component in the pathway (Choi et al. 1999).

SAS (shade avoidance syndrome) responses have been limiting factors in increasing crop yield at higher planting density. phyB, along with some of its interacting factors PIF4, PIF5 and PIF7, has been shown to be involved in the regulation of SAS responses in Arabidopsis (Smith and Whitelam 1997; Lorrain et al. 2008; Li et al. 2012). There have been several studies interested in modifying the phytochrome mediated light pathway for crop improvement. Overexpression of phytochromes (phyA and phyB) has been repeatedly tested to increase yield in densely-planted crops including rice, potato and tobacco, however it comes at the cost of unwanted pleiotropic

effects such as dwarfism and decrease in apical dominance (Robson et al. 1996; Thiele et al. 1999; Boccalandro et al. 2003; Schittenhelm et al. 2004; Garg et al. 2006; Rao et al. 2011). In order to efficiently manipulate the phytochrome mediated pathway in crops, especially monocots, a better understanding of the downstream components in the pathway is required. Hence, I was interested in identifying and cloning maize phytcohrome genes followed by identification of PhyB interacting factors.

Based upon the genome sequence and homology based analysis, the *Zea mays* cv. B73 genome has been predicted to contain 3 phytochrome genes with 2 homeologs each: *PhyA1* (NCBI Accession # AY234826), *PhyA2* (NCBI Accession # AY260865), *PhyB1* (NCBI Accession # AY234827), *PhyB2* (NCBI Accession # AY234828), *PhyC1* (NCBI Accession # AY234829) and *PhyC2* (NCBI Accession # AY234830). Monocots and dicots are believed to have split ~140-200 Mya and since then they have been evolving independently. Therefore, it has been suggested that the functions of monocot and dicot phytochromes may not be completely conserved (Sheehan et al. 2004). In fact, several studies have indicated unique roles of Arabidopsis and rice phytochromes (Takano et al. 2001; Franklin and Whitelam 2004; Takano et al. 2005). It therefore becomes even more important to identify and characterize phytochrome interacting factors in maize which could lead to information potentially useful for future biotechnological advancement of high yield varieties.

To identify phytochrome interacting factors, I screened a *Zea mays* cv. B73 etiolated seedling cDNA library using the yeast-two-hybrid system with the C-terminal domain of ZmphyB1 as bait. C-terminal domains of phytochromes instead of full length proteins have been used in the

past as bait since chromophore binding and photoconversion are not required for interaction to take place (Ni et al. 1998). I was expecting to obtain members of the PIF family as the primary interacting proteins as shown in Arabidopsis (Ni et al. 1998). Unlike Arabidopsis, maize contains two homeologs of each of three phytochromes. The presence of these duplicated phytochromes raises several questions including potentially diverged functions of Phys through differential ability to bind PIFs or other factors. Some of the PIFs and other phytochrome interacting proteins have been shown to interact with more than one phytochrome, especially phyA and phyB in Arabidopsis. It is also likely that interacting proteins identified for ZmphyB1-CTD would interact with other maize phytochromes as well. To test this, my approach was to amplify and clone other maize phytochromes and individually determine potential interactions with each maize phytochrome using targeted yeast-two-hybrid assay. Since transgenic crops overexpressing Arabidopsis phyA/phyB have been shown to have increased yield under high planting density, I was also interested to test possible cross species interaction between Arabidopsis and maize phytochrome-PIF components. For further confirmation of interactions, my goal was to establish an *in vitro* transcription-translation system for phytochromes and their interacting factors followed by reconstitution of maize holophytochromes by using phyocyanobilin (PCB). Subsequently, I wanted to determine if full-length maize phytochromes can photoconvert to Pr-Pfr forms under R and FR light, as indicated by potential binding of active Pfr form to the PIFs and other phytochrome interacting factors. With this study, I was able to find two PIF3 homeologs in maize and showed interaction of PIF3s with different maize and Arabidopsis phytochromes.

Materials and Methods

Yeast-two-hybrid screening

The C-terminal domain of ZmphyB1 was used as bait for the yeast two hybrid screening. Yeast strain YRG2 (Stratagene, La Jolla, Cat # 240060) competent cell preparation and transformation with GBD:*ZmPhyB1*-CTD construct was performed as per the standard protocol suggested by Frozen-EZ Yeast Transformation II Kit (Zymo Research, Irvine, CA, USA). Yeast-two-hybrid screening was performed as per Matchmaker Gold Yeast Two-Hybrid System (Clontech, CA; Cat # PT4084). An in-house size-fractionated cDNA library was used as screening prey, made from etiolated *Zea mays* cv. B73 seedlings using the Stratagene Lambda Hybri-Zap vector (Stratagene, La Jolla, CA). A positive control was created by co-transforming BD:*AtPHYB*-CTD and GAD:AtPIF3 constructs. Negative controls were created by co-transforming BD:Zm*PhyB1*-CTD and empty GAD vector (pGADT7 or pGAD424).

Colony lift filter assay:

A sterile precut Whatman filter paper was placed in a 150 mm Petri plate and soaked in freshly prepared Z-buffer-β-mercaptoethanol-X-gal solution (prepared as per Clontech Yeast Protocols Handbook). Another filter paper was placed on to the plate with colonies and gently pressed to transfer the colonies on to the filter paper. The filter paper was lifted off and dipped in liquid N_2 for 10-15 sec and thawed at room temperature 2-3 times. Then, the filter (colony side up) was

placed on top of the Z-buffer solution soaked filter paper. The plate was covered and incubated at 37˚C for 2-4 hours or more until blue color developed.

BLAST, sequence alignment and protein domain/motif analysis

BLASTs were performed by tools provided at NCBI, http://www.arabidopsis.org/ and http://www.maizesequence.org. Alignments were performed using the MAFFT & MUSCLE tools in Geneious Pro v5.6.5 (Biomatters Ltd., Auckland, New Zealand). For identification of domains and motifs in protein sequences, Pfam (http://pfam.sanger.ac.uk/) and prosite (http://prosite.expasy.org/) scan tools were used. Graphic representations of domains were created by DOG v2.0 (Ren et al. 2009).

RNA extraction from maize seedlings (for RT-PCR)

About 6 days old (after germination) maize seedlings were collected and immediately frozen in liquid N_2. The mass of seedling including root was about 0.32 g. The tissue was crushed in a pre-chilled mortar and pestle using liquid N_2. The crushed tissue (suspended in about 1-1.5 ml liquid N_2) was transferred to a pre-chilled 15 ml Corning conical bottom polypropylene tube. The powdered tissue was resuspended in 10 volumes of CTAB RNA extraction buffer by vortexing. The resuspension was sonicated for 90 sec. Equal volume of acidic phenol: chloroform: isoamyl alcohol (25:24:1) was added and vortexed for 1 min. The upper aqueous part was carefully removed and added to a pre-spun (1500xg for 2 min) 15ml heavy phase lock gel. Same amount

of chloroform: isoamyl alcohol was added and mixed several times by inverting. It was then centrifuged at 1500xg for 5 min. Aqueous phase was again carefully removed and transferred to a fresh 15 ml Corning tube. Then, 0.25 volume of 10M LiCl was added and RNA was let to precipitate overnight at 4˚C. The sample was divided into 1.5 ml X 3 microfuge tubes (1.5 ml) and processed separately. RNA was pelleted by centrifuging at 14k rpm at 4˚C for 30 min. Pellet was washed with pre-chilled 80% ethanol and centrifuged again at 14k rpm for 10 min. Pellet was resuspended in 225 µl of RNAse free water (Ambion, USA). 225 µl of ammonium acetate and 750 µl of ethanol was added and incubated at -20˚C for 1 hr. RNA was pelleted again by centrifuging at 14k rpm for 30 min. The pellet was vacuum-dried and resuspended in 50 µl RNase free water. The RNA was quantified using a UV-spectrophotometer and integrity was confirmed by agarose gel electrophoresis containing formaldehyde.

Reverse transcription

Reverse transcriptions were performed by using the standard protocol suggested by the manual for the Transcriptor First Strand cDNA synthesis kit (Roche Applied Science, IN, USA Cat # 04 379 012 001). *Zea mays* cv. B73 RNA (5 µg) was used as template with the following extension cycle: 65˚C for 45 min, 60˚C for 45 min, 85˚C for 5 min and 4˚C forever.

Big Dye terminator sequencing

The sequencing reactions were performed using BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Carlsbad, CA) with following setup: 200-300 ng template, 1X sequencing buffer, 10 pmol primer and 1 µl of Big Dye terminator v3.1 enzyme mix in a 20 µl setup with following program: 98˚C for 3 min, {98˚C for 15 sec, 48˚C for 20 sec, 60˚C for 4 min} for 35 cycles and 4˚C forever. Purification and further processing of sequencing reaction was performed by UIUC Core sequencing facility.

Molecular cloning of maize phytochromes and *Pif3*

ZmPhyB1

Maize *PhyB1* CDS (Maizesequence accession # GRMZM2G124532_T03) was amplified in two parts- N-terminal (1-2348) and C-terminal (1942-3534). N-terminal region was obtained by RT-PCR using MH289 (5'-ACGGGGCCGCCTAATCCAGA-3') and MH287 (5'-CAACTGTCGCCTCAGATTCCTTGAAGATAAGATCA-3') primers. The RT reaction was carried out as mentioned above with the MH287 (5'-CAACTGTCGCCTCAGATTCCTTGAAGATAAGATCA-3') reverse primer which represents around 2 kb region of *ZmPhyB1* CDS sequence. The amplified N-terminal fragment was subsequently cloned using TOPO TA cloning Kit (Invitrogen, CA, USA). The C-terminal fragment was amplified by RT-PCR using MH177 (5'-GATA**CATATG**CTTAGTTCCGTAGCAAGAGAGAT-3') and MH178 (5'-GAAA**GGATCC**GGGTCAGCACGGATCTTAAC-3') primers. NdeI and BamHI restriction sites

(marked as bold) were added in the forward and reverse primers respectively and were further used for subcloning the PCR product into the yeast plasmid vector pGBKT7 (Clontech Labs, CA, USA). The N-terminal and C-terminal fragments were joined by transformation assisted recombination method in yeast (Fig 1.6) (Larionov et al. 1996). The N-terminal fragment was re-amplified using the *ZmPhyB1*-NT-TOPO construct as template with following primers: ZPB-TAR-F (5'-ATGGAGGAGCAGAAGCTGATCTCAGAGGAGCATATGGCGTCGGGCAGCCGCGCCA-3') and MH-287 (5'-CAACTGTCGCCTCAGATTCCTTGAAGATAAGATCA-3'). *Pfu* Ultra HF DNA polymerase (2.5 U) (Stratagene, USA) was used with following parameters 94˚C for 2 min, {94˚C for 30 sec, 58˚C for 40 sec, 72˚C for 3 min}: 30 cycles, 72˚C for 10 min and 4˚C forever. The 5'-end of the amplicon contained a 31bp region which overlapped with the plasmid vector (pGBKT7 MCS sequence upstream of NdeI restriction site) and a 407bp overlap at the 5' end of C-terminal fragment. The *ZmPhyB1*-CT-pGBKT7 construct was linearized at the NdeI restriction site, the point across which the NT-fragment had to be inserted. The linearized plasmid and the NT fragment PCR (Purified by Qiagen PCR purification kit) were co-transformed in a molar ratio of 2.2:1 into yeast Y187 competent cells prepared as per Frozen-EZ Yeast Transformation II Kit (Zymo Research, Irvine, CA, USA). The transformation mix was spread onto SD-trp plates and incubated at 30˚C for 3-4 days. The positive yeast clones were screened by yeast colony PCR. Yeast colonies were resuspended in 50 µl zymolyase digestion buffer1 and 5 U of zymolyase (Zymo Research, CA, USA) and incubated at 30˚C for 30 min followed by 95˚C for 10 min. The yeast cell-suspension (5 µl) was used as template for PCR. PCR was performed with zPB 400bp Fwd (5'-GTCGCCTCACCACTCCGTTCCCT-3') and MH287 (5'-CAACTGTCGCCTCAGATTCCTTGAAGATAAGATCA-3') primers. Plasmid DNA was extracted from

the potential positive clones using Zymoprep Yeast Plasmid Miniprep I Kit (Zymo Research, CA, USA, Cat # D2001). The plasmid DNA extract of the positive yeast colonies were then transformed into One Shot TOP10 Chemical competent cells (Invitrogen, CA, USA, Cat # C4040-10). Plasmid DNA was extracted using Qiaprep Spin Miniprep Kit (Qiagen, CA, USA). Presence of the full length of the *ZmPhyB1* coding sequence (CDS) in the construct was confirmed by sequencing (UIUC Core Sequencing Facility).

Mutation correction of cloned *ZmPhyB1* CDS by Site directed mutagenesis

Three point mutations (C64T, G156T & A1883G) present in the pGBKT7-*ZmPhyB1* construct were fixed as per QuickChange Multi Site-Directed Mutagenesis Kit (Stratagene, CA, USA, Cat #200514). Following primers were used: zPB QC64 (5'-GCCCGAGGCGCCGCGTCACGCGCACC-3'), zPB QC156 (5'-CGGAGTCGGTCTCCAAGGCCGTCGCCCAGTACAC-3') and zPB QC1883 (5'-GCAGAGGGCACCAACAACTCAAAAGCCATTGTCAATGGACAAGTTC-3'). The nucleotide base correction was further confirmed by full length sequencing.

AtPHYB Y276H & *ZmPhyB1* Y274H

The *AtPHYB-Y276H* mutant construct was created by standard protocol suggested by QuickChange Multi Site-Directed Mutagenesis Kit (Stratagene, CA, USA, Cat #200514) using AtPB QC Y276-H (5'-GGTTATGATCGTGTTATGGTTCATAAGTTTCATGAAGATGAGC-3') primer and *ZmPhyB1-Y274H* mutant construct by zPB QC Y274-H (5'-GACCGTGTCATGGTGCACAGGTTCCATGAAG-3') primer. The pET3a-*AtPHYB* and pCR4-TOPO-*ZmPhyB1* constructs were used as template. Correct clones were screened by Big-dye terminator sequencing. *ZmPhyB1-Y274H* CDS was subcloned into pET17b plasmid expression

vector by standard digestion and ligation method under NdeI-KpnI restriction sites from pCR4-TOPO-*ZmPhyB1*-Y274H construct.

ZmPif3e

The extended version *ZmPif3e* CDS was amplified using *ZmPif3*-pADGAL4 construct as template which was obtained from *Zea mays* cv. B73 cDNA library (see Yeast-two-hybrid screening results). *Pfu* Ultra II Fusion HS DNA Polymerase (Stratagene, CA, USA, Cat # 600674) was used for PCR amplification along with zPIFIII Ext1 HindIII F (5'-AGCA**AAGCTT**CATGAGCAAGGAGCCCTGCTGCT-3') and MH256 (5'-CGAC**GAATTC**TCATGTTTCAGCCTCATTTCT-3') primers. PCR product was purified as per QIAquick PCR Purification Kit (Qiagen, CA, USA, Cat # 28104) and subcloned under HindIII and EcoRI restriction sites of pET17b-GAD plasmid vector using standard restriction digestion and ligation methods.

ZmPif3-P (ZmPif3e homeolog)

ZmPif3-P CDS was amplified from the maize cDNA library using GRMZM2G115960-T03-NdeI-F (5'-**CATATG**TCCGACAGCAACGACTTC-3') and GRMZM2G115960-T03-Rev (5'-GATTACTGGCGAAGATCTCTTCATC-3') primers and TAKARA Ex Taq DNA polymerase (Clontech Labs, CA, USA). The amplified CDS was cloned into pGEM-T easy plasmid vector (Promega, Madison, WI) as per the suggested protocol. The cloned CDS sequence was confirmed by Big dye terminator sequencing using T7, SP6 and 5960_T03_0.6F (5'-TGGAGAGGGATTGATGAACTTCTC-3') primers and following cycle: 95˚C for 1 min, {95˚C for 15 sec, 48˚C for 15 sec, 60˚C for 4.5 min}: 35 cycles, 4˚C forever. Four point mutations were found:

C228T (silent mutation), T485C, A893G and C1429T. The three mutations were fixed by the standard method suggested by QuickChange Multi Site-Directed Mutagenesis Kit (Stratagene, CA, USA, Cat #200514) using following primers: ZmPIF-5960-C485T-QC (5'-CCGCGCATACCAGTACCAGGTCCAGGC-3'), ZmPIF-5960-G893A-QC (5'-GACCAAAAAGGCTTGTGAGGTTGCAGTCGCTACTC-3') and ZmPIF-5960-T1429C-QC (5'-CCCAGGCACGCCACCACAAGGTCTTGG-3'). 5-alpha competent High efficiency *E. coli* cells (NEB, Ipswich, MA, Cat # C2987G) were used for final transformation. Fixed nucleotides were confirmed by sequencing as described above.

ZmPhyB2

Zm*PhyB2* CDS (Maizesequence accession # GRMZM2G092174_T01) was amplified in two parts: N-terminal fragment (1-2068 bp) and C-Terminal fragment (1976-3501 bp). The N-terminal fragment was amplified by RT-PCR using MH349 (zPB2-F) (5'-ATGGCGTCGGACAGTCGCCCCC-3') and MH354 (zPB2-2.07Rev) (5'-GGCCTGTCAACTCGGCAATCTTTGCATTCCAC-3') and TAKARA Ex Taq DNA polymerase (Clontech Labs, CA, USA) and cloned into pCR4-TOPO plasmid vector using TOPO TA Cloning Kit for Sequencing (Invitrogen, USA). Reverse transcription was performed as described above with MH350 (zPB2-R-RT) (5'-GGTGTCAGCACGGATCTTAACATATCAGAC-3'). The C-terminal fragment was amplified from maize cDNA library using MH384 (zPB2 1976F) (5'-CGATTGATAGAGACAGCAACAGTACCCATA-3') and zPB2 Sp 3'Rev (5'-TTAACATATCAGCTGATTTTCTCTACCAGCTGC-3') primers and cloned into pCR4-TOPO plasmid vector using TOPO TA Cloning Kit for Sequencing (Invitrogen, CA, USA). The C-terminal fragment was re-amplified using EcoRI-zPB2-1976F (5'-TAGA**GAATTC**CGATTGATAGAGACAGCAACAGT-3')

and zPB2-sp-BamHI-Rev (5’-TAGA**GGATCC**TTAACATATCAGCTGATTTTCTCTAC-3’) primers and *Pfu* Ultra II Fusion HS DNA Polymerase (Stratagene, CA, USA, Cat # 600674) and subsequently subcloned into EcoRI and BamHI restriction sites of pGBKT7 plasmid vector by standard digestion and ligation methods. The N-terminal fragment was fused with the C-terminal by transformation assisted recombination (TAR) method where the N-terminal region was amplified again with primers zPB2-TAR-F (5’-ATGGAGGAGCAGAAGCTGATCCTAGAGGAGCATATGGCGTCGGACAGTCGCCCCC-3’) and MH354 (zPB2-2.07 Rev) (5’-GGCCTGTCAACTCGGCAATCTTTGCATTCCAC-3’) using *Pfu* Ultra II Fusion HS DNA Polymerase. The 5’ end of the resulting amplicon contained a 36 bp region which overlapped with the plasmid vector (pGBKT7 MCS sequence upstream of NdeI restriction site) and a 93 bp overlap at the 5’ end of C-terminal fragment. zPB2-CT-pGBKT7 plasmid was linearized by EcoRI restriction digestion. The linearized C-terminal construct (190 ng) and purified N-terminal fragment PCR (150 ng) (molar ratio 3:1) were co-transformed in to AH109 yeast competent cells as described earlier and 200 µl of transformation mix was spread on to SD-trp plate and incubated at 30˚C for 4-5 days. Plasmid DNA was extracted from yeast colonies by Zymoprep Yeast Plasmid Miniprep I Kit (Zymo Research, CA, USA, Cat # D2001) and transformed into One Shot TOP10 Chemical competent cells (Invitrogen, CA, USA, Cat # C4040-10). Plasmid DNA was extracted using QIAprep Spin Miniprep Kit (Qiagen, CA, USA). Full length of *ZmPhyB2*-pGBKT7 CDS construct was confirmed by Big dye terminator sequencing (UIUC Core Sequencing Facility) which also revealed four point mutations: T250C, T2212C (silent), A2510G and A2944G. The full length *ZmPhyB2* CDS was reamplified from the pGBKT7 construct using AseI-zPB2-FL F (5’-CTATATTAATGGCGTCGGACAGTCGCCC-3’) and zPB2-Sp-3’-Rev (5’-

TTAACATATCAGCTGATTTTCTCTACCAGCTGCTA-3') primers and *Pfu* Ultra II Fusion HS DNA Polymerase and subcloned into pCR-Blunt II-TOPO plasmid vector using Zero Blunt TOPO PCR Cloning Kit (Invitrogen, CA, USA). The three mutations were fixed by the standard method suggested by QuickChange Multi Site-Directed Mutagenesis Kit (Stratagene, CA, USA, Cat #200514) using following primers:

ZmPhyB2 CDS was reamplified from pCR-Blunt II-TOPO-*ZmPhyB2* construct using AseI-zPB2-5'F and M13R-27 universal vector specific primer and *Pfu* Ultra II Fusion HS DNA Polymerase in order to introduce AseI restriction sites at 5' end of *ZmPhyB2* CDS sequence. A-overhangs were added to the PCR products by incubating purified PCR products with 1X Takara Ex Taq DNA Pol buffer, 0.25 mM dNTPs, 5 U Takara Ex Taq DNA Polymerase at 70˚C for 30 min. The PCR amplicon was then cloned into pGEM-T easy plasmid vector (Promega, Madison, WI) as per the suggested protocol.

ZmPhyA2

ZmPhyA2 CDS (Maizesequence accession # GRMZM2G181028_T01) was amplified in two parts: N-terminal fragment (1-1710 bp) and C-Terminal fragment (1517-3396 bp). The N-terminal fragment was amplified by RT-PCR using MH376 zPhyA-5F (5'- ATGTCTTCCTTGAGGCCTGCCCAG-3') and MH404 zPA1sp-1720 Rev (5'-TGGCAAACTCTTCATCTTGACAACCTCG-3') and TAKARA Ex Taq DNA polymerase (Clontech Labs, CA, USA) and cloned into pCR4-TOPO plasmid vector using TOPO TA Cloning Kit for Sequencing (Invitrogen, USA). The C-terminal fragment was amplified from maize cDNA library using MH403 zPA1sp-1517-F (5'-ACATGATCTGTGGAATGGCAGTGGCT`-3'), MH377 zPA1-3R (5'-TCAATGTCCAGCTGCTGAAGGAGCA -3') primers and TAKARA Ex Taq DNA

polymerase (Clontech Labs, CA, USA). The PCR product was cloned into pCR4-TOPO plasmid vector using TOPO TA Cloning Kit for Sequencing (Invitrogen, USA). The C-terminal fragment was re-amplified using zPA2-CT 1.6Kb NdeI-F (5'-CTATCATATGGGAGGTGCAAAGCATGATCC -3') and zPA2-CT BamHI-Rev (5'-CTATGGATCCTCATCGTCCAACTGCTGAAGGAG -3') primers and *Pfu* Ultra II Fusion HS DNA Polymerase (Stratagene, CA, USA, Cat # 600674) and subsequently subcloned into NdeI-BamHI restriction sites of pGBKT7 plasmid vector by standard digestion and ligation methods. The N-terminal fragment was fused with the C-terminal by transformation assisted recombination (TAR) method where the N-terminal region was amplified again with primers zPA2-NT-TAR-F (5'-CTGATCTCAGAGGAGGACCTGCATATGTCTTCCTCGAGGCCTGC-3') and zPA2-NT-1681-R (5'-GGAAAGCCTTAAAGGATAACCTAGGG-3') primers, *Pfu* Ultra II Fusion HS DNA Polymerase and following cycle: 94˚C for 3 min, {94˚C for 30 sec, 58˚C for 30 sec, 72˚C for 30 sec}(Boccalandro et al. 2003; Schittenhelm et al. 2004): 22 cycles, 72˚C for 3 min and 4˚C forever. The 5'-end of the resulting amplicon contained a 194 bp region which overlapped with the plasmid vector (pGBKT7 MCS sequence upstream of NdeI restriction site) and a 87 bp overlap at the 5' end of C-terminal fragment. zPA2-CT-pGBKT7 plasmid was linearized by NdeI restriction digestion. Linearized C-terminal construct and purified N-terminal fragment PCR (molar ratio 3:1) was co-transformed in to AH109 yeast competent cells as described earlier and 200 µl of transformation mix was spread on to SD-Trp plates and incubated at 30˚C for 4-5 days. Positive clones were screened by yeast colony PCR using T7 universal primer and zPA2-NT-1681R (5'-GGAAAGCCTTAAAGGATAACCTAGGG-3') as described earlier.

The two mutations were fixed by the standard method suggested by QuickChange Multi Site-Directed Mutagenesis Kit (Stratagene, CA, USA, Cat #200514) using following primers: zPA2-322TC-QC (5'-ATGCACCTGAAATGCTTACAACGGTCAGCC-3') and zPA2-2962TG-QC (5'-AGTCAAGTGCTGATAGGTTGCCAGGCTAAAG-3') on *ZmPhyA2*-pGBKT7 construct. In-house *E. coli* electrocompetent cells were used for the transformation step. 1 µl and 1.5 µl of DpnI treated PCR reaction was mixed to 50 µl of electrocompetent cells and incubated on ice for few minutes. Electric pulse at standard EcI settings was given to the mixture and 500 µl of prewarmed SOC medium was added to it. The transformation mix was shaken at 250 rpm at 37˚C for 1hr and then plated onto LB agar Kan^+ plates. Plasmid was extracted from few colonies and the nucleotide correction was confirmed via Big dye terminator sequencing.

ZmPhyC1

ZmPhyC1 CDS (Maizesequence accession # GRMZM2G057935_T01) was amplified in two parts: N-terminal fragment (1-1922 bp) and C-Terminal fragment (1828-3408 bp). The N-terminal fragment was amplified by RT-PCR using MH380 zPC1-5F (5'-ATGTCGTTGCCGTCGAACAACC-3') and MH383 zPC1-1923-R (5'-TTACCGGCAATGTCGACAGCCAA -3') and TAKARA Ex Taq DNA polymerase (Clontech Labs, CA, USA) and cloned into pCR4-TOPO plasmid vector using TOPO TA Cloning Kit for Sequencing (Invitrogen, USA). Reverse transcription was performed as described earlier with MH381. The C-terminal fragment was also amplified by RT-PCR using MH382 zPC1-1855-F (5'-CAGGGGCTACTTGAACTGAGAACAGTT-3') and MH381 zPC1-3R (5'-TCAGAATTTACTCGTCGAAGGCTTGGAC -3') primers and cloned into pCR4-TOPO plasmid vector using TOPO TA Cloning Kit for Sequencing (Invitrogen, USA). The overlapping region (1828-1922

bp) between the N-terminal and C-terminal fragments contain a unique restriction site SalI. The C-terminal fragment was cut out from the pCR4-TOPO-*ZmPhyC1*-CT construct by SalI and NotI (present near 3'-end of insert in the plasmid vector) restriction digestion and ligated at the same sites in pCR4-TOPO-*ZmPhyC1*-NT to obtain the full length CDS. Full length sequence was confirmed by Big Dye terminator sequencing which revealed a mutation a G->T mutation at position 512. The mutation was fixed by QuickChange Multi Site-Directed Mutagenesis Kit (Stratagene, CA) as per the suggested protocol using zPC1 QC 512TG (5'-ACGCCATATTGCACC**G**GATCGACGTCGGCC-3') primer. *ZmPhyC1* was reamplified from the pCR4-TOPO construct using zPC1-5P-NdeI-F (5'-GATC**CATATG**TCGTTGCCGTCGAACAACC-3'), T7 universal primer (vector specific) and *Pfu* Ultra II Fusion HS DNA Polymerase in order to introduce NdeI and EcoRI restriction sites at 5' and 3' ends respectively. A-overhangs were added to the PCR products by incubating the purified PCR products with 1X Takara Ex Taq DNA Pol buffer, 0.25 mM dNTPs, 5 U Takara Ex Taq DNA Polymerase (Clontech Labs, CA, USA) at 70˚C for 30 min. The PCR amplicon was then cloned into pGEM-T easy plasmid vector (Promega, Madison, WI) as per the suggested protocol.

Yeast two hybrid constructs

Bait constructs (DNA-BD:Phy translational fusion) were created by subcloning C-terminal regions of all the phytochrome CDS into pGBKT7 plasmid vector (Clontech Labs, CA, USA). *ZmPhyB1*-CT (1933-3534 bp) was amplified using zPB1-NdeI-1933-F (5'-CTAG**CATATG**ATAAATGAGCTTAGTTCCGTAG-3'), *ZmPhyB2*-CT (1936-3501bp) with ZmB2-AseI-

1937-F (5’-GATT**ATTAAT**GGGATAAATGAGCTTAGCTCTG-3’), *ZmPhyA2*-CT (1831-3410bp) with zPA2-NdeI-1831-F (5’-TCGA**CATATG**CTTGATGGGCTTGCTGAATTGC-3’) and *ZmPhyC1*-CT (1831-3435bp) with zPC1-NdeI-1831-F (5’- CGAT**CATATG**GGGCTACTTGAACTGAGAACAG-3’) as forward primers. Templates used for PCR were the corresponding *Phy*-pET17b plasmid constructs. Plasmid vector specific T7-Terminator (5’-GCTAGTTATTGCTCAGCGGT-3’ *Primer sequence not exactly same as T7 Terminator universal primer and only works for pET17b plasmid vector) was used as reverse primer for all due to which some restriction sites from the plasmid vector pET17b were also included at the 3’ end of the amplicons. *ZmPhyB1*-CT and *ZmPhyA2*-CT were subcloned under NdeI-BamHI and *ZmPhyC1*-CT under NdeI-EcoRI restriction sites of pGBKT7 plasmid vector. *ZmPhyB2*-CT was subcloned under NdeI-BamHI restriction sites of plasmid vector pGBKT7 however AseI-BamHI restriction sites were used in the insert as *ZmPhyB2* CDS contains an NdeI site at position 3114. AseI and NdeI restriction sites generate compatible cohesive ends. Both AseI and NdeI sites were lost after ligation.

Prey constructs (AD:PIF translational fusion) were created by subcloning *ZmPif3e* and *ZmPif3-P* CDS into pGADT7 plasmid vector (Clontech Labs, CA, USA). *ZmPif3-P* was subcloned by restriction digesting *ZmPif3-P*-pGEM-T-easy construct and ligating the eluted fragment into pGADT7 using NdeI-EcoRI restriction sites.

3-AT titration

The BD yeast transformants were streaked onto SD-trp-his agar plates along with 0, 5, 10, 15, 20 and 25 mM 3-Amino-1,2,4-triazole (3-AT) in order to titrate the minimum amount of 3-AT

required to inhibit leaky histidine expression for each. The plates were incubated at 30˚C for until 10-12 days and colony growth was monitored. The minimum concentration of 3-AT at which no growth was observed was selected (25 mM for *ZmPhyB2* and 10mM for all other constructs, Fig 2.12) for further experiments.

Targeted yeast two hybrid

Yeast AH109 competent cells were prepared and transformed with 0.5-1 µg plasmid DNA of binding domain (BD) containing constructs: *ZmPhyB1*-CT-pGBKT7, *ZmPhyB2*-CT-pGBKT7, *ZmPhyA2*-CT-pGBKT7, *ZmPhyC1*-CT-pGBKT7, *AtPHYB*-CT-pGBT9 and pGBKT7 using protocol suggested by Frozen-EZ Yeast Transformation II Kit (Zymo Research, Irvine, CA, USA). Transformants were grown onto SD-trp agar plates at 30˚C for 4-5 days or more. These transformant colonies were picked up and again competent cells were made. BD-construct containing yeast competent cells were each transformed with 0.5-1 µg plasmid DNA of activation domain (AD) constructs: *ZmPif3e*-pGADT7, *ZmPif3-P*-pGADT7, *AtPIF3*-pGAD424 and pGADT7 using the same kit. 75-100 µl of transformation mix was spread onto SD-leu-trp agar plates and incubated at 30˚C for 4-5 days or more. 4-5 colonies from each transformation were pooled and grown into 3 ml of SD-Leu-Trp broth at 30˚C, 250 rpm for 1-2 days until good growth was observed. A loop-full of each grown culture was streaked onto SD-Leu-Trp, SD-Leu-Trp-His+3-AT and SD-Leu-Trp-His-Ade+3AT agar plates and incubated at 30˚C for about 10-12 days while monitoring growth. SD-Leu-Trp-Ade+3AT agar plate grown colonies were further taken to β-gal colony lift assay as described earlier.

Immunoprecipitation assay constructs

For immunoprecipitations, GAD:*ZmPIF3e* and GAD:*ZmPIF3-P* translation fusion constructs (GAL4 activation domain fused in frame with the N-terminus of full length PIF3e or PIF3-P) were created in pET17b expression vector (Novagen (Merck KKaA), Madison, WI) where expression is under the control of T7 promoter. pET17b-GAD vector was created by amplifying GAD sequence from pAD-GAL4-2.1 vector and cloned into pET17b vector under NdeI-HindIII restriction sites. *ZmPIF3e* CDS was subcloned into pET17b-GAD plasmid vector as described earlier. *ZmPif3-P* was PCR amplified with HindIII_ZmPIF3-P-5'F (5'- ATAT**AAGCTT**CATGTCCGACAGCAACGAC-3') and ZmPIF3-P-EcoRI-Rev (5'- AGGC**GAATTC**TATTTTTGTAGTATTTGTG-3') primers, *Pfu* Ultra II Fusion HS DNA Polymerase and following cycle: 94˚C for 3 min, {94˚C for 30 sec, 58˚C for 30 sec, 72˚C for 30 sec}: 26 cycles, 72˚C for 3 min and 4˚C forever. The amplified PCR fragments were digested with HindIII-EcoRI and ligated into pET17b-GAD plasmid vector. Construct sequence was confirmed by Big dye terminator sequencing as described earlier. Full lengths of phytochrome CDS were subcloned into the pET17b vector. *ZmPhyB1* CDS was subcloned under NdeI-KpnI, *ZmPhyA2* under NdeI-BamHI; *ZmPhyC1* under NdeI-EcoRI and *ZmPhyB2* under AseI/NdeI-BamHI of pET17b plasmid vector. *ZmPhyA2* was subcloned into pET17b plasmid vector by digesting out from *ZmPhyA2*-pGBKT7 construct and ligating under NdeI-BamHI restriction sites. *ZmPhyC1* CDS was subcloned into pET17b plasmid vector by sequencing digesting pGEM-T-easy-*ZmPhyC1* construct with BsaI (to linearize the construct) followed by NdeI-EcoRI and further ligating into NdeI-EcoRI of pET17b plasmid vector. *ZmPhyB2*-pET17b construct was created by cutting out *ZmPhyB2* from XmnI linearized pGEM-T-easy-*ZmPhyB2* construct by AseI-EcoRI restriction digestion followed by ligating into pET17b plasmid vector

under NdeI-EcoRI sites (AseI and NdeI restriction sites generate compatible ends). Sequencing of the pET17b- *ZmPhyB2* construct revealed a new mutation T2212C which was fixed by QuickChange Multi Site-Directed Mutagenesis Kit (Stratagene, CA) as per the suggested protocol using zPB2-QC-2212-CT (5'-AGCTGAAGACATTTGGG**T**CAGAGCAATCTAAGGGTG-3') primer. 5-alpha competent High efficiency *E. coli* cells (NEB, Ipswich, MA, Cat # C2987G) were used for final transformation. GAD:*AtPIF3* and *AtPHYB* constructs were used as described previously (Ni et al. 1998; Ni et al. 1999). All of the constructs were verified by Big dye terminator sequencing.

Immunoprecipitation assay

Bait and prey proteins were synthesized *in vitro* using ^{35}S-Methionine in the TNT T7 Quick Coupled Transcription/Translation System (Promega, Madison, WI; Cat # L1170). For each immunoprecipitation, 15 µl of Protein A magnetic dynabeads (Invitrogen, CA, USA) was used. 30 µl of beads were washed with 0.1ml of 1X binding buffer (PBS buffer pH 7.25, 0.1% Tergitol NP-40 (Sigma), 1 mM EDTA, 0.1 mg/ml BSA, 1 c0mplete mini protease inhibitor (Roche Applied Science, IN, USA) per 10 ml of PBS buffer). Antibody binding was performed by incubating beads with 100 µl of 1X binding buffer with ~3 µg of Anti-GAD antibodies (Santacruz Biotechnology, CA, Cat # 1663) and rotating at 4˚C for 1.5 hr. Supernatant was removed by using Dynal MPC magnetic stand (Invitrogen, CA) and beads were resuspended in 24 µl of bait TNT and 96µl of 1x binding buffer (1:5 ratio) and again rotated at 4˚C for 3-4 hours. Phytochrome TNT was precleared and holophytochrome was prepared by adding 5 µl of

prewashed protein A-dynabeads and 40 µM of bacterial Phycocyanobilin (Frontier Scientific, Logan, UT, Cat # P14137) to 95 µl of phytochrome TNT and 95 µl of 2X binding buffer and rotating at 4˚C for 1 hr in dark. Using magnetic stand, holophytochrome was collected and divided ~ 95 µl each into 2 1.5ml microfuge tubes. Both tubes were incubated under highest fluence of red light for 10 min after placing them horizontally onto ice. One tube (Pfr) was removed and stored in cool dark place and the other tube was treated with 5 min of far-red (Pr). Bait beads were gently resuspended and divided equally into two 1.5 ml microfuge tubes. Bait beads were captured and supernatant was removed by using magnetic stand. Pr and Pfr holophytochromes were added to one of each bait-bead containing tube, mixed gently and rotated at 4˚C for 4 hr in dark. Pull down was ended was collecting beads with the magnetic stands and washing the beads once with 0.5 ml 1x binding buffer and twice with 0.5 ml 1X wash buffer (binding buffer without BSA). Before the last wash, the beads were transferred to a new tube. Bead bound proteins were eluted by resuspending beads into 30 µl of 2X Laemmli's buffer (with DTT) and heating at 65 ˚C for 5 min. The protein samples were run on 4-20% gradient SDS PAGE (Bio-Rad, Hercules, CA). The gel was fixed in acetic acid:methanol:glycerol (7:7:10% solution) by shaking for 10 min and then dried in Bio-Rad 583 gel dryer for 1.5-2 hr. The dried gel was then exposed to phosphor screen overnight and scanned using Storm 840 Phosphorimager (Amersham). All light sensitive steps were performed in dark room under green safe light. The spectrum of green safe light, red light and far red light used is shown in Fig 2.18 measured using Stellarnet EPP200 spectroradiometer.

Results

Identification of phytochrome interacting proteins

Previously, it was shown that the C-terminal domain of phytochromes is capable of interacting with the signaling partners and has been used to screen for identifying interacting proteins via the yeast two hybrid system (Ni et al. 1998; Ni et al. 1999). Based upon a similar strategy, I used the C-terminal domain of maize PHYB1 (647-1161 aa) as bait in the yeast two hybrid system and screened a maize cDNA library (Fig2.2). After screening, His and β-galactosidase positive colonies were selected (Fig 2.3A) and the corresponding prey plasmids were rescued and further analyzed. Sequencing and BLAST analysis of the prey plasmids revealed coding sequences for a bHLH protein (Maizesequence Transcript Id # GRMZM2G387528_T02 & Genbank accession # ACR35132) homologous to Arabidopsis PIF3 (TAIR accession # AT1G09530.1) and this gene was referred to subsequently as *ZmPIF3* (Fig 2.4). Another potential interacting protein was found to be of unknown identity, however it contained a BTB/POZ (Bric-a-brac, Tramtrack, Broad-complex/ Pox virus and Zinc finger) domain belonging to the NPH3 (Non-phototropic hypocotyl 3) superfamily. The corresponding homologs in Arabidopsis are NPY family of proteins (Tair accession # AT2G14820; AT5G67440; AT4G37590) shown to possess a role in auxin regulated organogenesis (Cheng et al. 2008).

Extended ORF of maize *Pif3* gene

As per Maizesequence Transcript Id # GRMZM2G387528_T02 & Genbank accession # ACR35132, the *ZmPif3* ORF consists of 1917 nt coding for 638 amino acids. However, upon careful analysis of the *ZmPif3* cDNA sequence obtained from the yeast two hybrid screening, an additional translational start site in the same ORF was found 93 nt upstream of the one reported which would correspond to additional 31 amino acids at the N-terminal end (Fig 2.5). A similar region was also found to be present at the N-terminal end of AtPIF3 protein (Fig 2.6). The extended version of *ZmPIF3* has been referred to as *ZmPIF3e*. In an immunoprecipitation study, GAD:ZmPIF3 did not show binding to Pfr form of ZmphyB1 (Fig 2.19), however GAD:ZmPIF3e showed interaction with the ZmphyB1 (Pfr) (Fig 2.20 B) indicating that the extended 31 amino acid region in ZmPIF3e is required for interaction. ZmPIF3e was used for further studies.

Molecular cloning of maize phytochromes and *Pif3* CDS

Based upon the predicted coding sequences of maize *PhyA1* (Genbank accession # AY234826), *PhyA2* (AY260865), *PhyB1* (AY234827), *PhyB2* (AY234828), *PhyC1* (AY234829) and *PhyC2* (AY234830) (Sheehan et al. 2004), I cloned full-length ORFs for *PhyA2*, *PhyB1*, *PhyB2* and *PhyC1* from maize cDNA libraries and/or by RT-PCR. The *ZmPhyA2* ORF contains 3396 nt coding for 1131 amino acids; *ZmPhyB1* contains 3486 nt coding for 1161 amino acids; *ZmPhyB2* contains 3501 nt coding for 1166 amino acids; and *ZmPhyC1* ORF contains 3408 nt coding for 1135 amino acids. Full-length PCR amplifications of phytochrome CDS from *Zea mays* cv. B73 cDNA

library or RT-PCR were not successful, probably because of uneven GC distribution (higher towards N-terminal and lower towards C-terminal ends). Hence, the CDSs were amplified into two halves and joined together by the transformation assisted recombination (TAR) cloning method (Fig 1.6). *ZmPif3e* CDS was amplified from the cDNA containing prey vector obtained from the yeast-two -hybrid screening. *ZmPif3e* ORF contains 2010 nt and encodes 669 amino acids. The maize genome also contains a *ZmPif3* paralog (homeolog) located on chromosome 3 (Maizesequence accession # GRMZM2G115960). We refer to this gene as *ZmPif3-P*. The complete ORF of *ZmPif3-P* contains 1704 nt encoding 567 amino acids. The *ZmPif3-P* CDS was subsequently cloned from the maize cDNA library used for the two-hybrid study.

Domain analysis of maize phytochromes and PIF3s

Typically, plant phytochromes can be divided into several domains: the N-terminal PAS (Per, Arn and Sim), GAF (cGMP-stimulated phosphodiesterase) and PHY domains and C-terminal two consecutive PAS and an HKRD (Histidine kinase related domain) (Bae and Choi 2008). All the six maize phytochromes show similar highly conserved domain structure and location (Fig 2.7). The homeologs of PHYA (PHYA1 and PHYA2) and PHYC (PHYC1 and PHYC2) are identical in length with 1131 and 1135 amino acids respectively. In contrast, PHYB2 (1166 aa) is slightly longer than PHYB1 (1161 aa). The two homeologs of each maize phytochrome show very high levels of similarity to each other with PHYA1 and PHYA2 (96%), PHYB1 and PHYB2 (94%) and PHYC1 and PHYC2 being 94% identical. In Arabidopsis, the stronger binding affinity of phyB to PIF3 compared to phyA has been attributed to an additional 37 amino acid extension at the N-

terminus of phyB (Fig 2.8A) (Zhu et al. 2000). Compared to ZmPHYA1/A2, ZmPHYB1/B2 also possess an extended region of 20 amino acids at the N-terminus; however there is a second extended region of 6-10 residues rich in glycine and alanine located 12 residues downstream of the first one (Fig 2.8B). This second extended region is not very prominent in AtPHYB. As shown in Fig 2.8C, the sequences in the extended regions of ZmPHYB1/ZmPHYB2 are not strongly conserved with AtPHYB. The significance of the two extended regions in ZmPHYB1/B2 remains to be determined experimentally.

In Arabidopsis, PIF3 contains a conserved bHLH domain, an APB (active phytochrome B binding) motif and APA (active phytochrome A binding) motif. APB and APA motifs have been shown to be essential for binding to phyB and phyA respectively (Khanna et al. 2004; Al-Sady et al. 2006). ZmPIF3e and the paralog ZmPIF3-P retain a similar domain structure with an APB domain at the N-terminal end, APA in the middle region and bHLH in the C-terminal region (Fig 2.9 & 2.10). ZmPIF3e is significantly larger (669 aa) compared to ZmPIF3-P (567 aa) which is close to the length of AtPIF3 (524 aa). The bHLH domain is highly conserved between the AtPIF3 and both ZmPIF3 paralogs. APB/APA motif regions also show similarity between Arabidopsis and maize PIF3. However, there is a low level of conservation outside these domains. The two maize paralogs PIF3e and PIF3-P also show a high level of similarity across the gene with 69% identity (Fig 2.11).

PIF3e interacts with C-terminal domain of phyB1, phyB2 and phyA2 in maize

In Arabidopsis, PIF3 has been shown to be involved in both phyA and phyB mediated signaling pathways and both the C-terminal and N-terminal domains of phyA and phyB have been shown to interact with PIF3 (Ni et al. 1998). As the structure of the maize phytochrome family is different compared to dicots and it has been suggested that monocot and dicot phytochrome functions may not be totally conserved (Sheehan et al. 2004), I identified other members of the maize phytochrome family which can potentially interact with ZmPIF3e. Using the GBD:Phy-CTD translational fusion of *ZmPhyA2*, *ZmPhyB1*, *ZmPhyB2* and *ZmPhyC1* as bait and GAD:*ZmPIF3e* as prey in a targeted yeast two hybrid assay, growth of yeast co-transformant colonies in -His and -Ade selective plates was observed for *ZmPhyA2* (Fig 2.15A-B), *ZmPhyB1* (Fig 2.13A-B) and *ZmPhyB2* (Fig 2.14A-B) and confirmed for all these genes by colony filter lift assay for β-galactosidase expression (Fig 2.15C; 2.13C; 2.14C). There was very slow growth even in the non-selective liquid culture medium for all the co-transformants which included *ZmPhyC1* and PIFs and almost no growth at all on SD medium plates even after 10-12 days of plating indicating a potential toxicity of ZmPHYC1 to yeast cells (Fig 2.16).

PIF3-P also interacts with C-terminal domain of maize phyB1 and phyA2

The maize genome contains two paralogs (homeologs) of the Arabidopsis PIF3 gene. It is possible that differential binding of phytochromes with the two paralogs acts as a regulatory mechanism of light signal pathway in maize. Co-transformed yeast colonies with *PhyA2* or *PhyB1* together with *PIF3-P* showed growth on -His and -Ade selective media, indicating that

the C-terminal domains of maize phyA2 and phyB1 also interact with PIF3-P (Fig 2.15A-B; Fig 2.13A-B). An observation of blue color on a colony filter lift β-galactosidase assay (Fig 2.15C; Fig 2.13C) also provided further confirmation. Colonies with *ZmPhyB2* and *ZmPif3-P* showed slow growth on non-selective medium and little to no growth on selective media (Fig 2.14A-B). *ZmPhyC1* and *ZmPif3-P* cotransformants did not grow on non-selective or selective media at all (Fig 2.16B). Hence, the data on binding potential of ZmphyB2 and ZmphyC1 to ZmPIF3-P remains inconclusive and further work is required.

Maize and Arabidopsis phytochromes and PIF3 cross-interact

Although Arabidopsis and maize are quite evolutionarily distant organisms, their phytochromes show a high level of homology including their domain structure. Similarly, Arabidopsis and maize PIF3s possess a high level of conservation in the bHLH domain and APB/APA motifs (Fig 2.11). While there is little conservation outside of the bHLH and APB/APA domains, it could still be sufficient for similar function and potential cross-interaction between Arabidopsis and maize PHY-PIF3. Results from my targeted yeast two hybrid assay confirmed that C-terminal domains of maize phyA2, phyB1 and phyB2 also interact with AtPIF3 (Fig 2.13; 2.14; 2.15). Not surprisingly, in the reverse experiment, the C-terminal domain of AtphyB showed interactions with both ZmPIF3e and ZmPIF3-P (Fig 2.17) indicating a conserved function at least in terms of phyA/phyB binding to PIF3.

In vitro maize holophytochrome reconstitution and interaction with PIF3

In Arabidopsis, it has been shown that both N-terminal and C-terminal regions of phytochromes possess binding determinants for PIF3 and they act synergistically in full length phytochrome (Ni et al. 1999). PIF3 binding was observed in holophytochrome (phytochrome apoprotein + chromophore) only in the active Pfr form formed under red light but not with apoprotein or the inactive Pr form (Ni et al. 1999). Arabidopsis phytochrome has been demonstrated to utilize bacterial phycocyanobilin (PCB), which is structurally similar to phytochromobilin (PφB), and reconstitute functional holophytochrome (Parks and Quail 1991; Kunkel et al. 1995; Ni et al. 1999; Kami et al. 2004). As maize phytochromes show high levels of homology to Arabidopsis phytochromes, it is likely that maize phytochromes would also be capable of autocatalytically binding to PCB and thus forming holophytochrome. To test this, I performed an *in vitro* immunoprecipitation assay using GAD:ZmPIF3e as bait and Pr/Pfr forms of the four maize phytochromes i.e. phyA2, phyB1, phyB2 and phyC1, produced from clones described earlier in this chapter using *in vitro* transcription, translation and chromophore assembly, as prey. Results show that ZmphyB1 was indeed able to make holophytochrome and photoconvert between the Pfr and Pr forms in response to R and FR light, indicated by the preferential binding of the Pfr form to ZmPIF3e but not Pr form (Fig 2.20 B). The other three maize phytochromes, phyA2, phyB2 and phyC1 that were tested in this assay did not show any binding activity, as indicated by the absence of corresponding phytochrome in the pellet (Fig 2.20 C-D).

ZmphyB1 interaction with NPH3L protein

Yeast two hybrid screening of the maize cDNA library using ZmphyB1-CTD as bait showed another potential interacting protein encoded by the gene Maizesequence accession # GRMZM2G172506. A BLAST against Arabidopsis proteins indicated that the product of this gene encodes a protein similar to Arabidopsis NPY2 (NAKED PINS IN YUC MUTANTS 2) or NPH3 like (Non-phototropic hypocotyl 3) protein (TAIR accession # AT2G14820) and hence, we named it as ZmNPH3L. The NPY2/NPH3 family of proteins has been shown to be involved in auxin mediated organogenesis, phototropism and root gravitropism (Furutani et al. 2007; Li et al. 2011). I wanted to check if ZmNPH3L protein also interacts with full length ZmphyB1 in a light dependent manner. I performed an immunoprecipitation assay using GAD:ZmNPH3L as bait and ZmphyB1 in Pr and Pfr forms as prey. As shown in Fig 2.21, ZmNPH3L binding was not observed in either Pr or Pfr form of the phytochrome.

Discussion

Recently, maize (*Zea mays* cv. B73) genome was sequenced (Maize Genome Sequencing Project) which greatly facilitated the study of the components of phytochrome mediated pathway at the molecular level. The maize genome contains 3 phytochrome lineages *PhyA*-*PhyC* with two homeologs of each, and the homeologs have been suffixed with 1 and 2 (Mathews and Sharrock 1996). Proteins encoded by all the six maize phytochrome genes show highly conserved domain structure having the GAF, PAS, PHY and HKRD domains that are found in Arabidopsis (Fig 2.7). Maize PHYA1-A2 and PHYC1-C2 pairs are identical in length (1131 & 1135 amino acids respectively), however PHYB1 & B2 differ slightly from each other (1161 and 1166 amino acids respectively). Similar to Arabidopsis, maize PHYB1/PHYB2 proteins are also longer compared to PHYA1/PHYA2 showing a presence of additional 20 amino acid region at N-terminal end (Fig 2.8). Similar extended region in AtPHYB has been attributed to higher binding affinity of AtphyB to AtPIF3 (Zhu et al. 2000). In addition to the 20 amino acid stretch of the extended region, there is also a second extended region of 6-10 residues rich in glycine and alanine located further downstream in ZmPHYB1/B2. This second extended region is not very prominent in AtPHYB. The sequence in the extended regions of ZmPHYB1/B2 is not well conserved relative to corresponding region in AtPHYB. However, the AtPHYB extended region is rich in glycine with at least two GGGR repeats. This is similar to the second extended region of ZmPHYB2 which contains a stretch of glycines followed by three alanine residues. The second extended region of ZmPHYB1 is shorter and contains fewer glycines compared to ZmPHYB2. It is

possible that the glycine-rich region acts as a determinant for selective binding or providing additional binding affinity of ZmphyB1/B2 for its interacting partners. In that case, the second extended region of ZmPHYB1/B2 would more likely be associated with these features. It is also possible that the varying length of the second extended region in ZmPHYB1 and ZmPHYB2 provides a different function or binding affinity that is yet to be determined.

Yeast two hybrid screening revealed two potential maize phyB1 interacting proteins: a bHLH domain containing transcription factor (Maizesequence accession # GRMZM2G387528_T02) and a BTB/POZ domain containing NPY2/NPH3 family protein. The bHLH transcription factor showed homology to AtPIF3, which was also the first PIF family member identified in Arabidopsis. It was hence named as ZmPIF3. Upon further analysis, the ZmPIF3 cDNA sequence was found to have an additional in frame translational start site relative to the maize genome annotation. The extended sequence region was also significantly conserved in the AtPIF3. It was determined that the extended region was necessary for interaction with phyB1. The extended version was named as ZmPIF3e and was used for all further analysis.

The maize genome also contains a homeolog of the ZmPIF3e gene (Maizesequence accession # GRMZM2G115960) which we refer to as ZmPIF3-P. The existence of the two PIF3 homeologs in maize warranted the testing of their interaction with other phytochromes. In addition to the conserved bHLH domain, both ZmPIF3e and ZmPIF3-P also contain APB (Active Phytochrome B binding) motifs and APA (Active Phytochrome A binding) motifs. APB and APA motifs have been shown to be essential for binding of PIF3 to phyB and phyA respectively in Arabidopsis. Presence of APB and APA motifs indicated that both PIF3e and PIF3-P can potentially interact

with ZmphyB1/B2 and ZmphyA1/A2. In order to test these findings, I amplified and cloned *ZmPhyA2*, *ZmPhyB1*, *ZmPhyB2* and *ZmPhyC1* CDS from maize cDNA library/RT-PCR. Amplification of *ZmPhyA1* and *ZmPhyC2* from mRNA was not successful, which could be due to absence of these transcripts from the cDNA library and the seedling RNA as a result of differential expression of homeologs. In a targeted yeast two hybrid system using the C-terminal domains of each cloned maize phytochrome as bait, I was able to show that ZmPHYA2 interacts with both of the PIF homeologs (Table 3.1). ZmPHYB1-CTD also showed interaction with both maize PIF3e and PIF3-P. ZmPHYB2-CTD did show binding to ZmPIF3e, and the interaction test with ZmPIF3-P remained inconclusive because of inhibited growth of co-transformed yeast cells. ZmPHYB2-ZmPIF3-P interaction can be further tested with an immunoprecipitation assay. *ZmPhyC1* yeast transformant cells were also very slow in growth possibly indicating a lack of interaction with ZmPIF3e, although the slow growth could be partly attributed to yeast toxicity of ZmPHYC1, making the experiment harder to interpret. *ZmPhyC1-ZmPif3-P* co-transformants showed no growth at all. Again, immunoprecipitation assay would be a better approach to confirm any possible interaction as gene toxicity is likely to be a problem in these transformants.

Maize and Arabidopsis are far apart evolutionarily, thus it was not clear whether phytochromes and PIFs from these two organisms can cross-interact. By using a targeted yeast two hybrid system, I was able to demonstrate that CTD of ZmPHYA2, ZmPHYB1 and ZmPHYB2 can interact with AtPIF3. In another test, AtPHYB-CTD also showed interaction with both ZmPIF3e and ZmPIF3-P. The results for ZmPIF3 were also confirmed using immunoprecipitation. These results suggest that phytochrome-PIF3 binding and function between Arabidopsis and maize are quite

conserved, and also that the three dimensional interaction surfaces of PIF3 and phyotchromes remain fully compatible despite the evolutionary distance and significant differences in amino acid primary sequence.

The interactions observed using the C-terminal domains of phytochromes need not necessarily extrapolate to "full length" phytochrome because of the involvement of multiple domains, the need for chromophore assembly and light mediated conformational changes. By using protein produced in an *in vitro* transcription and translation system, I reconstituted holophytochrome by adding phycocyanobilin, a structurally similar chromophore found in cyanobacteria. Full-length holophytochromes were photoconverted by red light and red/far-red light to create the Pfr and Pr forms respectively. Similar to the results from Arabidopsis (Ni et al., 1999), the Pfr form of ZmphyB1 bound to ZmPIF3e but not the Pr form. This indicates a conformational change between the Pr and Pfr form of maize phyB1 consistent with that observed previously in Arabidopsis. Surprisingly, maize phyB2 did not bind to PIF3e, suggesting preferential binding of ZmPIF3e with the ZmphyB1 homeolog. It is possible that ZmphyB2-ZmPIF3e binding is merely weaker compared to that of ZmphyB1-ZmPIF3e, and that this weaker binding was not observed under the experimental conditions used in immunoprecipitation. Further experiments with lower stringency such as low salt and/or low concentration of detergent in the IP buffer, reduced number of washes could be performed to verify this, but it is evidence for differential specificity of ZmPIF3e for ZmphyB1 over ZmphyB2. It is possible that ZmphyB2 specifically binds to the PIF3-P homeolog, a hypothesis that still needs to be tested. Another possibility is that ZmPhyB2 does not interact with ZmPIF3 at all, but may perhaps act as a regulatory element of

phyB mediated signaling by making heterodimers with ZmphyB1 potentially reducing the binding properties of the dimer.

In Arabidopsis, both the C-terminal domain and full length phyA proteins were found to interact with PIF3. My targeted yeast two hybrid results also indicated that the C-terminal portion of one of the PHYA homeologs in maize, PHYA2, interacts with both PIF3e and PIF3-P. However, I was not able to identify any binding of ZmPIF3e with full length phyA2 holophytochrome (Pfr or Pr) in an immunoprecipitation assay. As observed in the case of maize phyB1 and phyB2, it is again possible that only phyA1 interacts with PIF3e and not phyA2. Unfortunately, I was not able to clone *PHYA1* CDS to test this hypothesis, as it does not appear to be expressed in seedlings. In Arabidopsis, phyA-PIF3 binding was also reported to be weaker than phyB-PIF3 (Zhu et al. 2000). Since the assay conditions used are quite stringent, experiments with lower stringency conditions may reveal a weaker interaction.

Expressed ZmPHYC1-CTD was likely toxic to the yeast cells, the absence of yeast growth on selective plates can't be interpreted as non-interaction with certainty. It is possible that other yeast strains like YRG2 or Y187 might work better for this construct. Immunoprecipitation assay between ZmphyC1 and ZmPIF3e also did not show binding. To confirm, more immunoprecipitation experiments need to be performed.

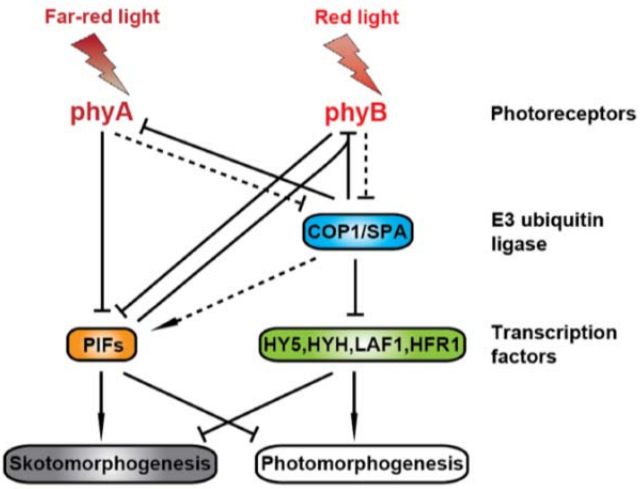
It is possible that the duplication of all three phytochrome genes in maize has led to selection for inactive versions of these genes, as the higher gene dosage may produce maladaptive phenotypes. This would be consistent with our observation that one homolog of each of phyA

and phyC are not expressed at detectable levels in seedlings, and that while both phyB homologs are expressed, one phyB homolog is inactive in PIF3 binding. Such negative selection against one copy of a dosage-dependent gene is not unknown, but severe selection pressure must be active to have inactivated one of all three phytochrome homeologous pairs within a relatively short evolutionary timeframe.

Another yeast two hybrid potential ZmphyB1 interacting protein, ZmNPH3L, did not show any binding to the Pr or Pfr forms of ZmphyB1 in immunoprecipitation assays where GAD:ZmNPH3L was used as bait. It is possible that ZmNPH3L requires post translational modification to interact with phytochrome which probably was not available in the *in vitro* system I used. This modification might not have been required to interact with only the C-terminal domain of ZmphyB1. It is also possible that the fusion of GAD domain to ZmNPH3L interferes the interaction with full-length ZmphyB1.

In conclusion, this study suggests that the biochemical function and molecular structure of the phy-PIF3 system is highly conserved between Arabidopsis and maize, despite sequence differences.

Figures and Tables



(Li et al. 2011)

Fig 2.1: A simplified model of the phytochrome signaling pathway showing the probable positions of various downstream components of phyA and phyB mediated pathways. The PIF family of bHLH transcription factors is primary intermediates in the pathway.

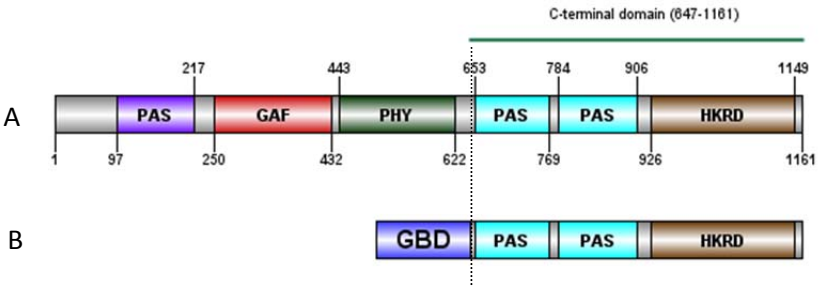
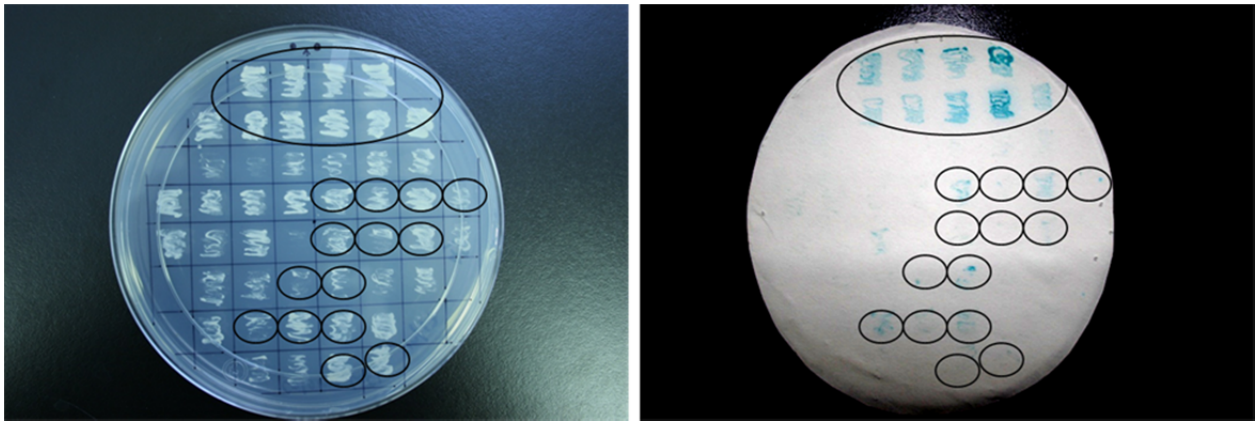


Fig 2.2: Domain structure of maize PHYB1. Schematic representation of (A) full length ZmPHYB1 terminal PAS (Per, Arn and Sim), GAF (cGMP-stimulated phosphodiesterase), PHY (Phytochrome) and HKRD (His Kinase Related domain) domains and (B) the C-terminal region with attached GBD (GAL4 DNA binding domain) used as bait in yeast two hybrid screening.

A



B

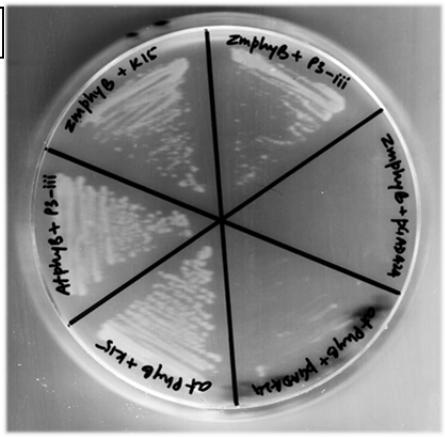


Fig 2.3: Yeast two hybrid screening using ZmPHYB1-CTD as bait: (A) Left- Potential His-positive Colonies Right- Colony lift assay showing the β-galactosidase positive colonies. Marked 23 colonies were selected for further analysis. (B) Further interaction confirmation by targeted yeast two hybrid showing cross-interaction between ZmPHYB with AtPIF3 and AtPHYB with ZmPIF3. [K15: AtPIF3; P3-iii: ZmPIF3; pGAD424: empty AD vector]

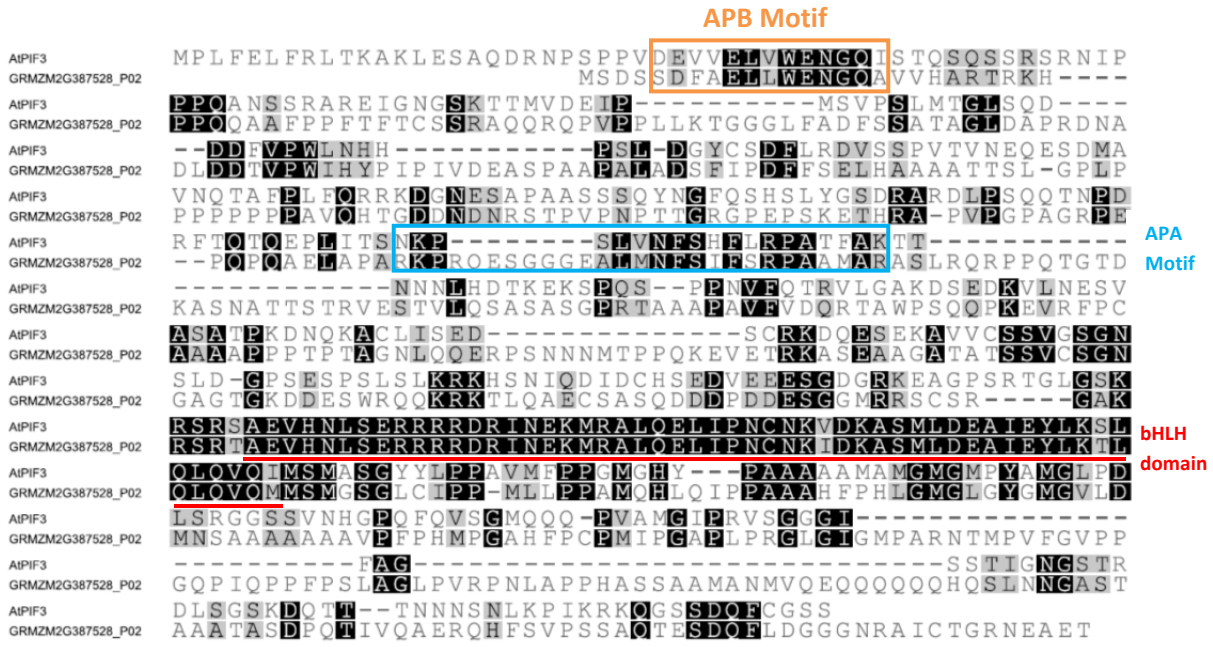


Fig 2.4: An alignment between AtPIF3 and the homolog in maize identified through yeast two hybrid screening. APB (Active Phytochrome B binding) motif, APA (Active Phytochrome A binding) motif and the highly conserved bHLH domain (underlined in red) are typical features of PIF3. There is low level of conservation outside of these domains/motifs.

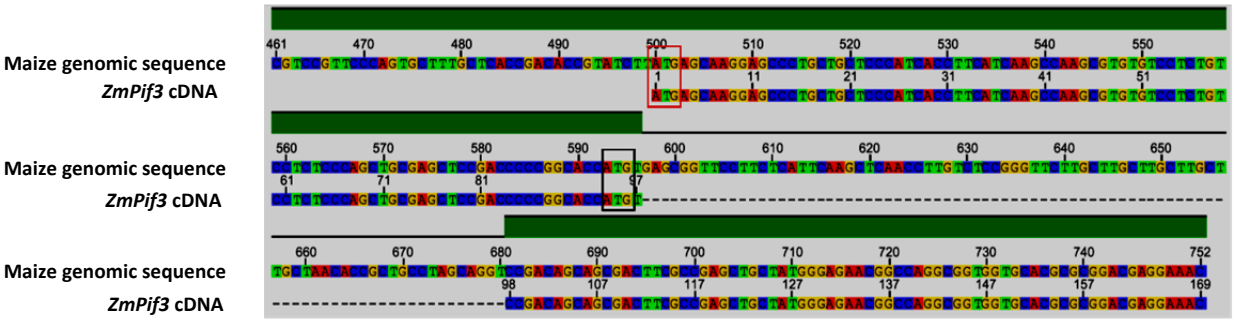
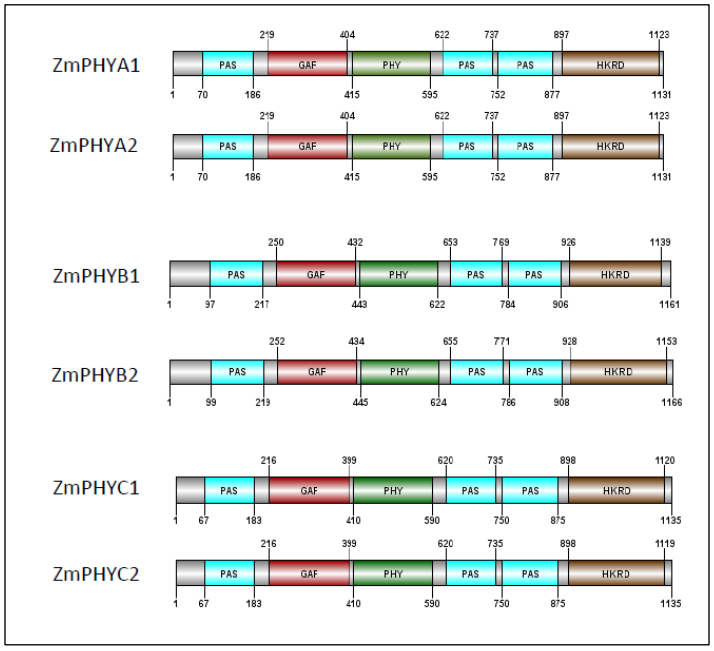


Fig 2.5: Sequence alignment of the N-terminal region of *ZmPif3* cDNA and corresponding genomic DNA showing the translational start site identified in the maize genome assembly and annotation (black box) and the in frame start site present at 93nt upstream (red box).

Fig 2.6: Alignment of the N-terminal region of AtPIF3 and the extended ZmPIF3 protein sequences. Originally predicted translational start site marked with (*). Underlined sequence represents the extended region in ZmPIF3e protein.



Phytochromes	% Identity	
	Protein	DNA
A1-A2	95	93
B1-B2	93	94
C1-C2	93	94

PHYA-B-C: ~50%

Fig 2.7: Schematic representation of domain structure of all maize phytochromes showing the highly conserved PAS, GAF, PHY and HKRD domain locations (Left Panel). Percent identity between the homeologs and the three maize phytochromes (Right Panel).

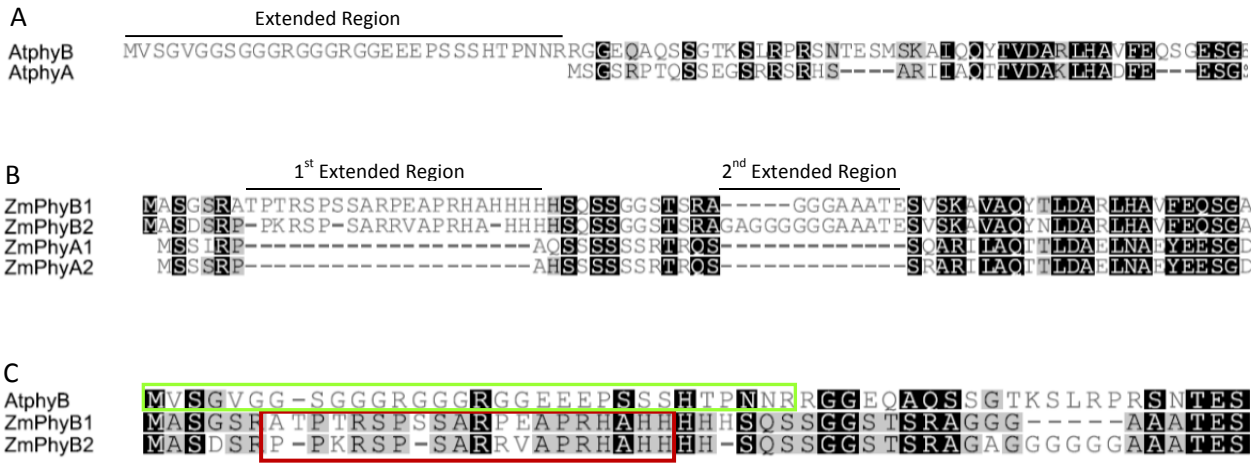


Fig 2.8: Alignment showing extended region at N-terminal end of PHYB compared to PHYA (A) Arabidopsis PHYB vs. PHYA (B) Maize PHYB1/B2 vs. PHYA1/A2. The 2nd extended region is relatively shorter and rich with glycine and alanine residues (C) Comparison of the N-terminal extended region in AtPHYB (green box) vs. ZmPHYB1/ZmPHYB2 (red box).

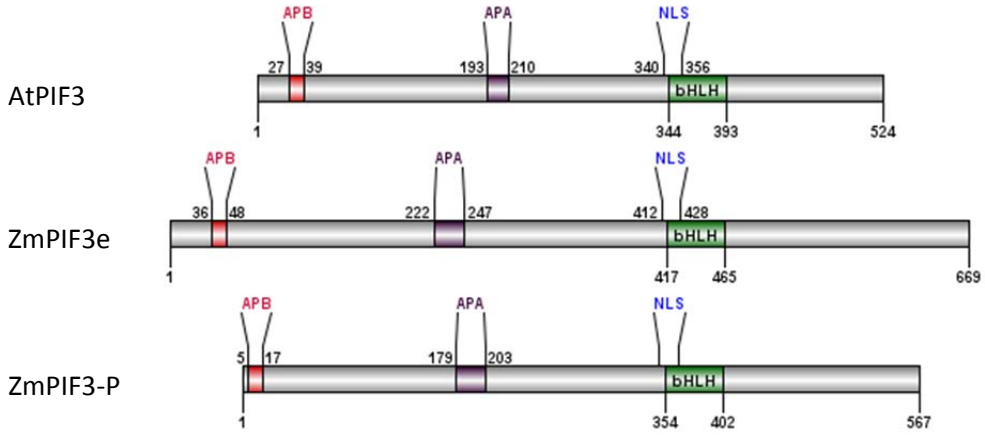


Fig 2.9: Schematic diagram of Arabidopsis PIF3, Maize PIF3e and PIF3-P showing the relative positions of bHLH domain, APB (Active Phytochrome B binding motif) and APA (Active Phytochrome A binding motif).

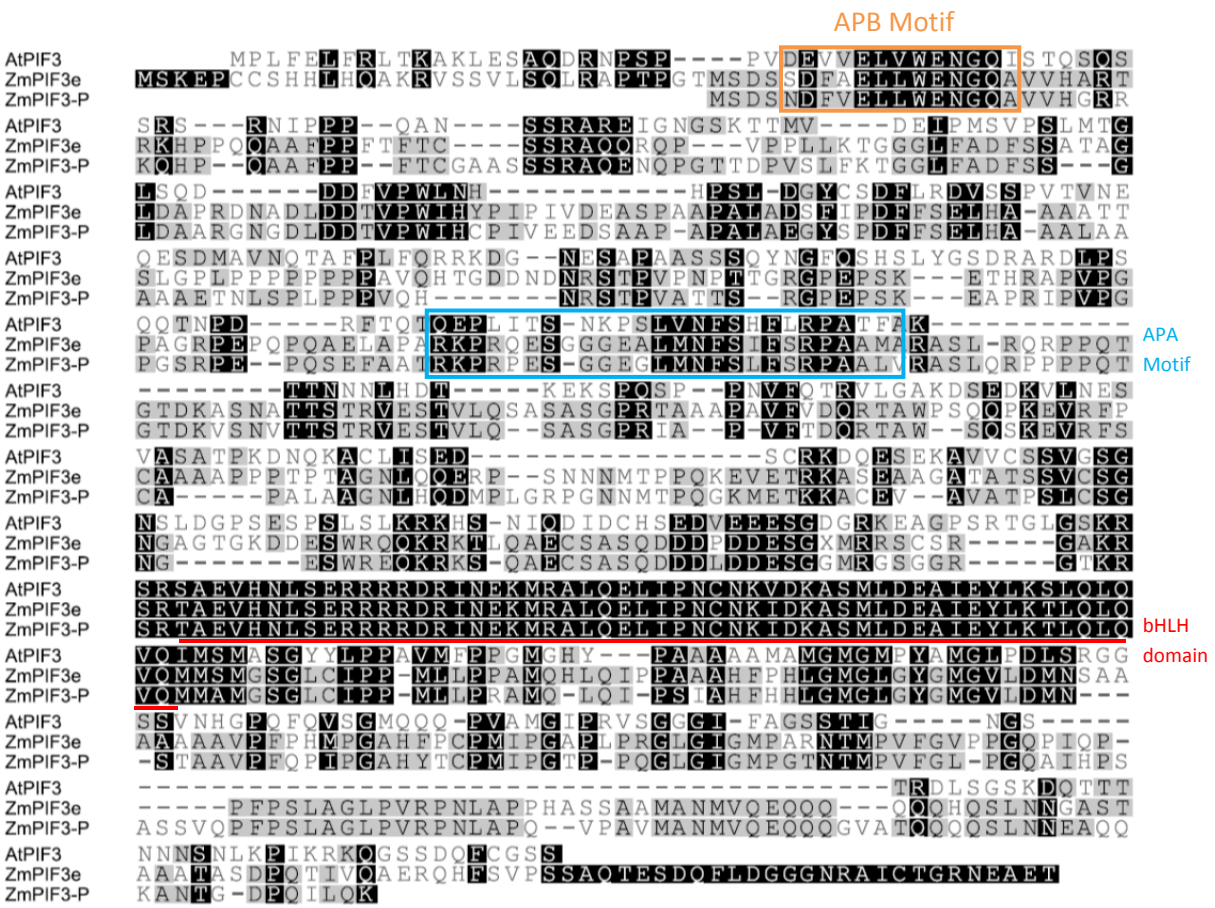


Fig 2.10: Alignment showing the conserved APB, APA and bHLH (red underlined) domains in AtPIF3 and maize counterparts. Note the low level of conservation outside these domains between Arabidopsis and maize.

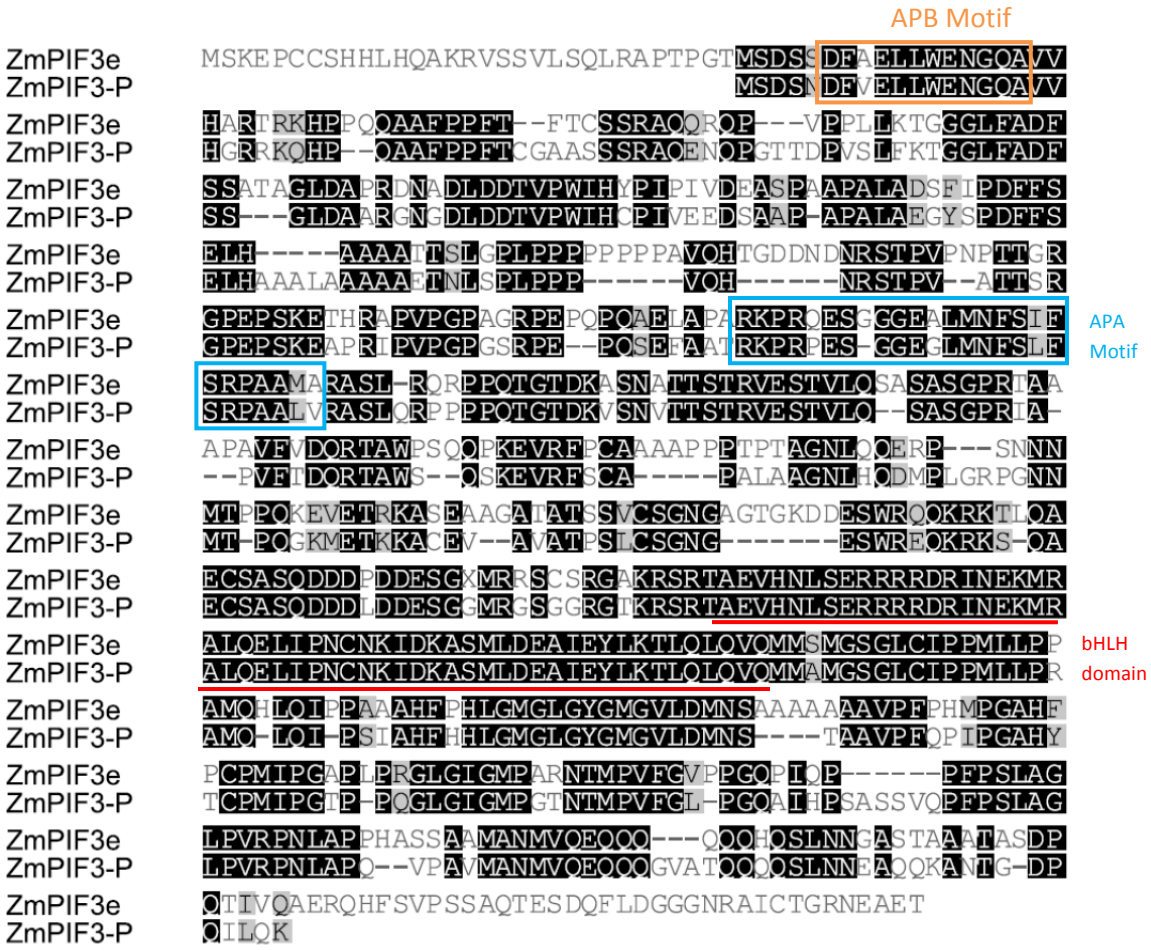


Fig 2.11: Pairwise alignment of the two maize PIF3 paralogs, ZmPIF3e and ZmPIF3-P showing high level of similarity and position of conserved APB/APA motifs and bHLH domain. ZmPIF3e protein is longer at both the N-terminal and C-terminal ends compared to that of ZmPIF3-P.

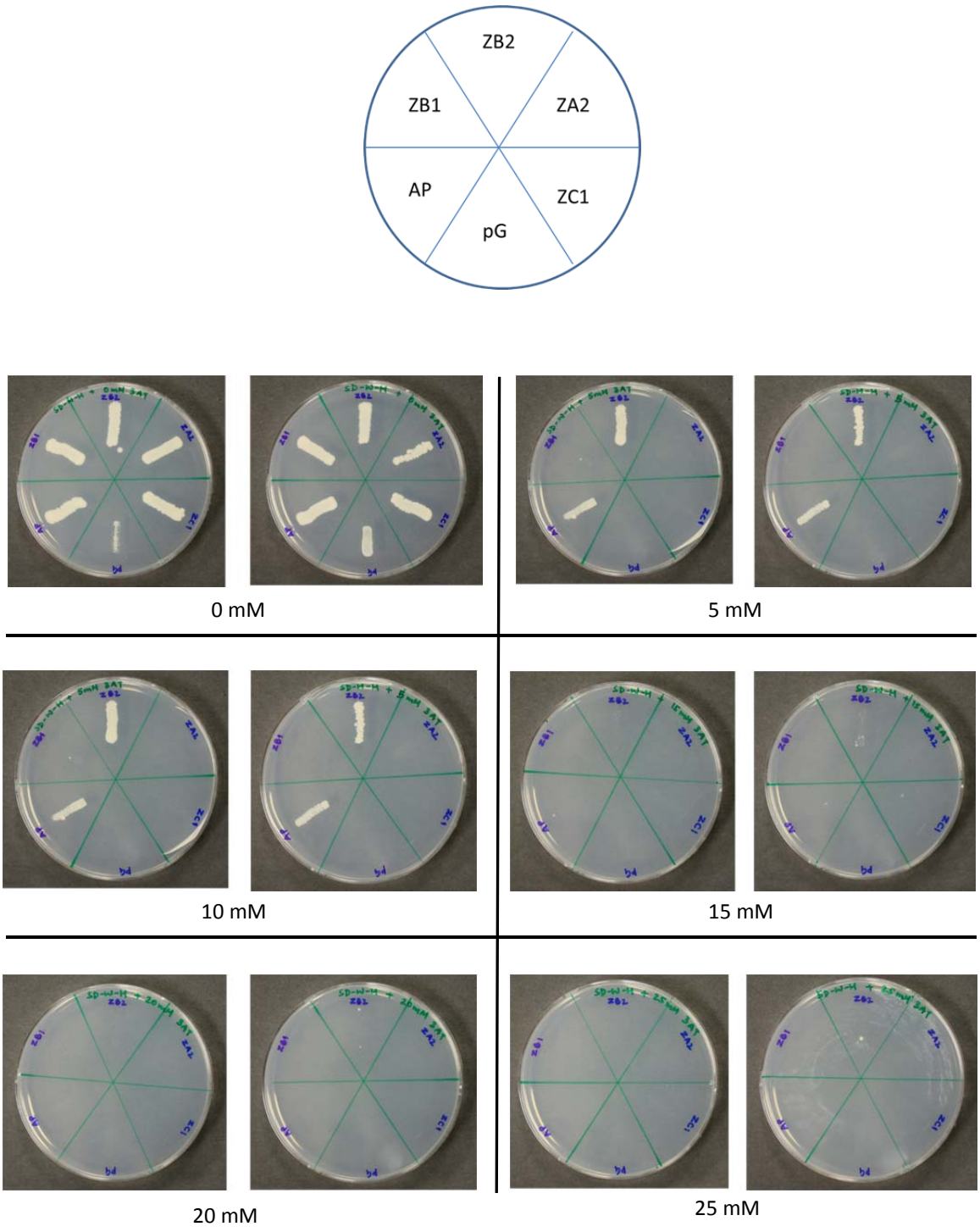


Fig 2.12: 3-AT titration for various GDB: phytochrome C-terminal domain constructs. Top: Plating plan. Transformed yeast cells were grown on SD-W-H. [ZB2: *ZmPhyB2*; ZA2: *ZmPhyA2*; ZC1: *ZmPhyC1*; pG: pGBKT7 vector only; AP- *AtPHYB*; ZB1- *ZmPhyB1*]. 25mM for ZmPhyB2 and 10mM 3-AT for all other constructs was selected for targeted yeast two hybrid assays.

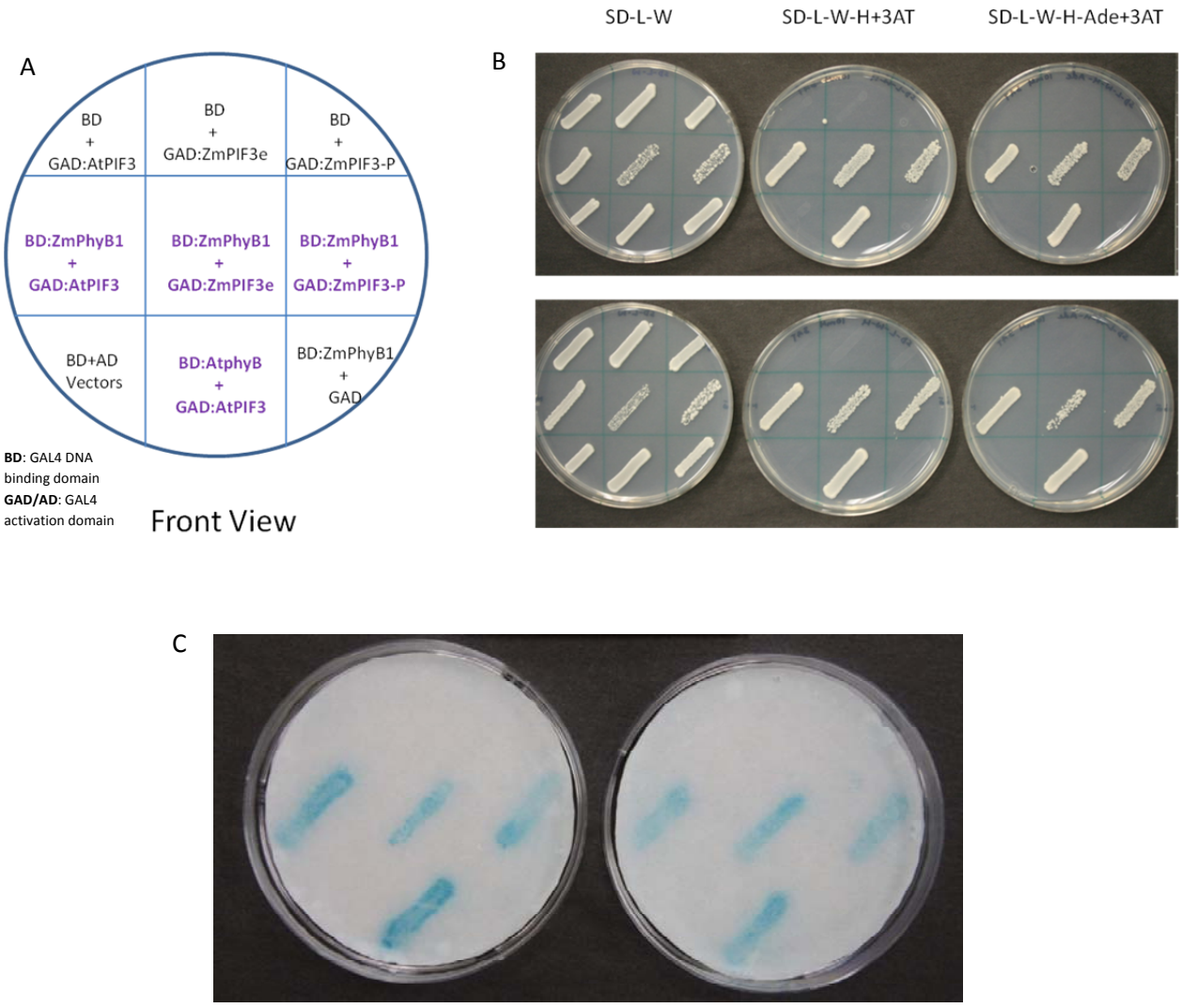


Fig 2.13: Targeted yeast two hybrid showing interaction between ZmPHYB1 and ZmPIF3e/ZmPIF3-P/AtPIF3 (A) Plating plan: Top row- empty bait vector negative controls; middle row- bait and prey constructs (magenta); Bottom row- empty bait+prey vector (BD+AD), positive control and bait with empty prey vector in the same order (B) Co-transformant colonies on non-selective (-L-W) and selective growth plates -L-W-H + 10mM 3-AT and –L-W-H-Ade+10mM 3-AT. Growth on selective growth plates indicates potential interaction. (C) Colony lift filter assay for β-galactosidase expression confirming the interaction. [SD: synthetic dropout medium; L: Leucine; W: tryptophan; H: histidine; Ade: adenine; 3AT: 3-amino-1,2,4-triazole]

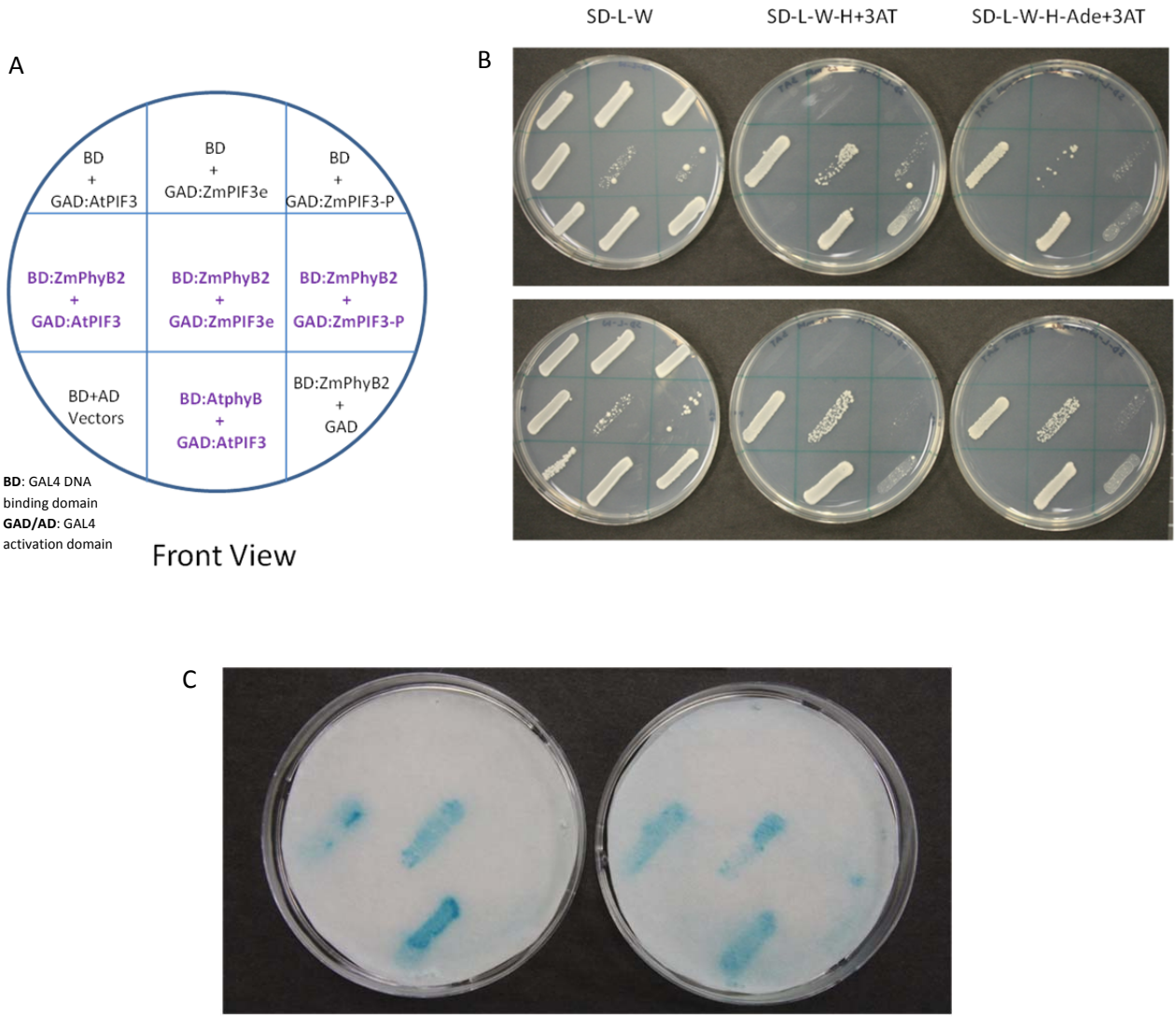


Fig 2.14: Targeted yeast two hybrid showing interaction between ZmPHYB2 and ZmPIF3e/ AtPIF3 (A) Plating plan: Top row- empty bait vector negative controls; middle row- bait and prey constructs (magenta); Bottom row- empty bait+prey vector (BD+AD), positive control and bait with empty prey vector in the same order (B) Co-transformant colonies on non-selective (-L-W) and selective growth plates -L-W-H+25mM 3-AT and –L-W-H-Ade+25mM 3-AT. Growth on selective growth plates indicates potential interaction. (C) Colony lift filter assay for β-galactosidase expression confirming the interaction. [SD: synthetic dropout medium; L: Leucine; W: tryptophan; H: histidine; Ade: adenine; 3AT: 3-amino-1,2,4-triazole]

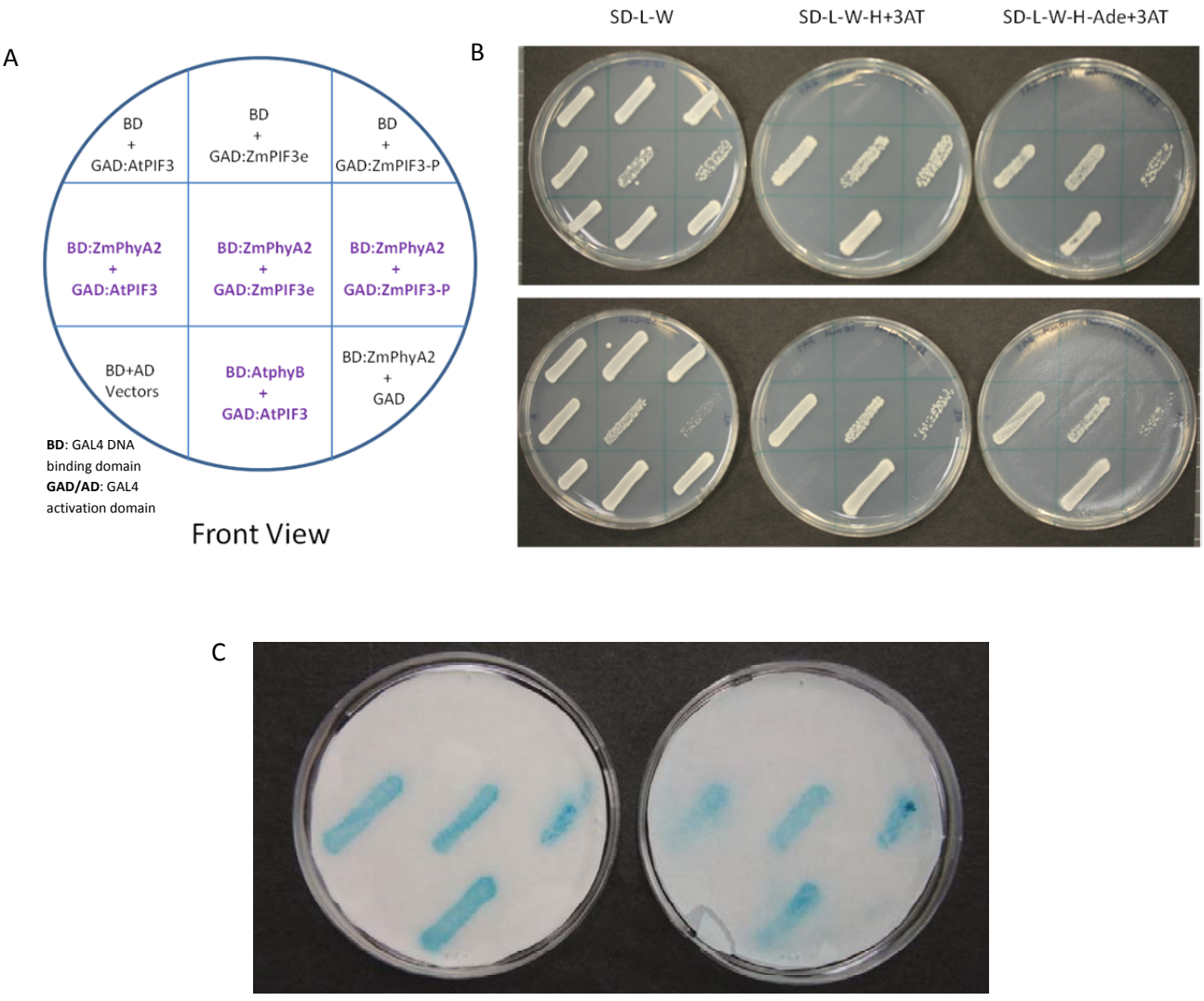


Fig 2.15: Targeted yeast two hybrid showing interaction between ZmPHYA2 and ZmPIF3e/ZmPIF3-P/AtPIF3 (A) Plating plan: Top row- empty bait vector negative controls; middle row- bait and prey constructs (magenta); Bottom row- empty bait+prey vector (BD+AD), positive control and bait with empty prey vector in the same order (B) Co-transformant colonies on non-selective (-L-W) and selective growth plates -L-W-H+10mM 3-AT and –L-W-H-Ade+10mM 3-AT. Growth on selective growth plates indicates potential interaction. (C) Colony lift filter assay for β-galactosidase expression confirming the interaction. [SD: synthetic dropout medium; L: leucine; W: tryptophan; H: histidine; Ade: adenine; 3AT: 3-amino-1,2,4-triazole]

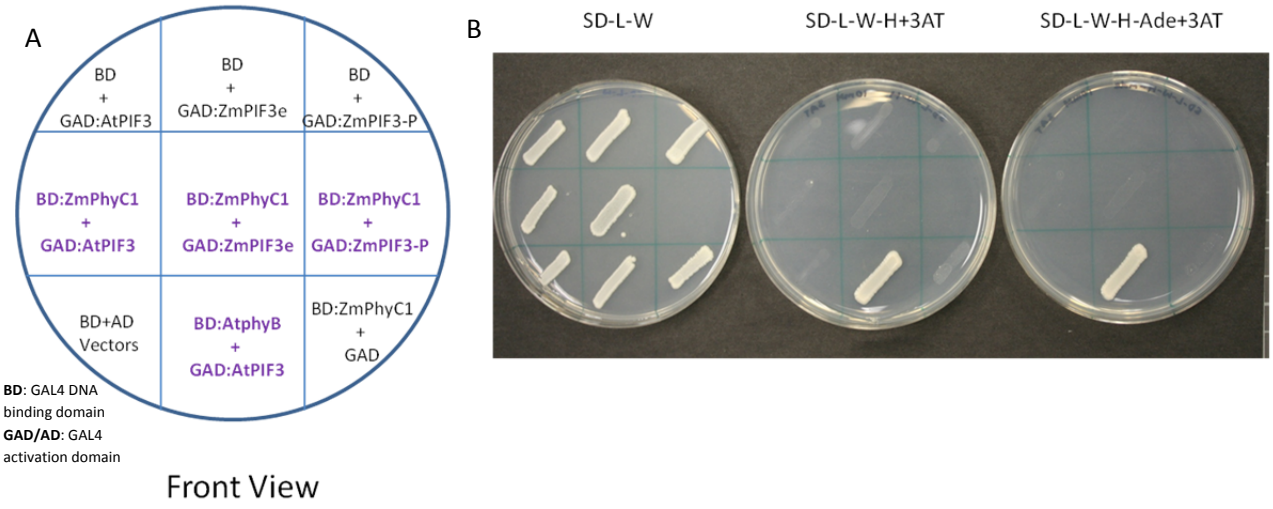


Fig 2.16: Targeted yeast two hybrid to check potential interaction between ZmPHYC1 and AtPIF3/ZmPIF3e/ZmPIF3-P (A) Plating plan: Top row- empty bait vector negative controls; middle row- bait and prey constructs (magenta); Bottom row- empty bait+prey vector (BD+AD), positive control (AtPHYB+AtPIF3) and bait with empty prey vector in the same order (B) Co-transformant colonies on non-selective (-L-W) and selective growth plates -L-W-H+10mM 3-AT and –L-W-H-Ade+10mM 3-AT. Growth on selective growth plates indicates potential interaction. Note no growth in selective growth plates except for the positive control [SD: synthetic dropout medium; L: Leucine; W: tryptophan; H: histidine; Ade: adenine; 3AT: 3-amino-1,2,4-triazole]

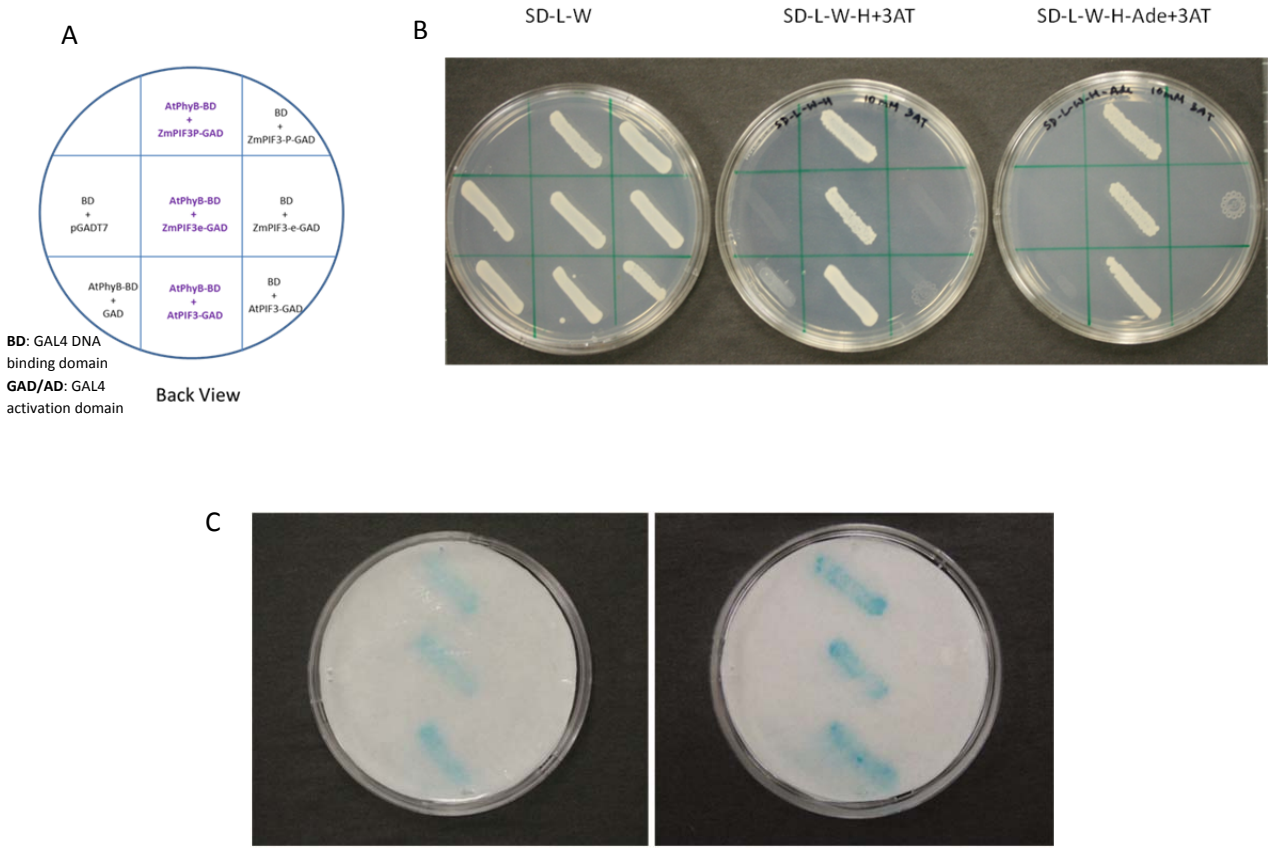


Fig 2.17: Targeted yeast two hybrid showing interaction between AtPHYB and ZmPIF3e/ZmPIF3-P (A) Plating plan: Top row- empty bait vector (BD) negative controls; middle row- bait and prey constructs (magenta); Bottom row- empty bait+prey vector (BD+AD), positive control and bait with empty prey vector in the same order (B) Co-transformant colonies on non-selective (-L-W) and selective growth plates -L-W-H+10mM 3-AT and –L-W-H-Ade+10mM 3-AT. Growth on selective growth plates indicates potential interaction. (C) Colony lift filter assay for β-galactosidase expression confirming the interaction. [SD: synthetic dropout medium; L: Leucine; W: tryptophan; H: histidine; Ade: adenine; 3AT: 3-amino-1,2,4-triazole]

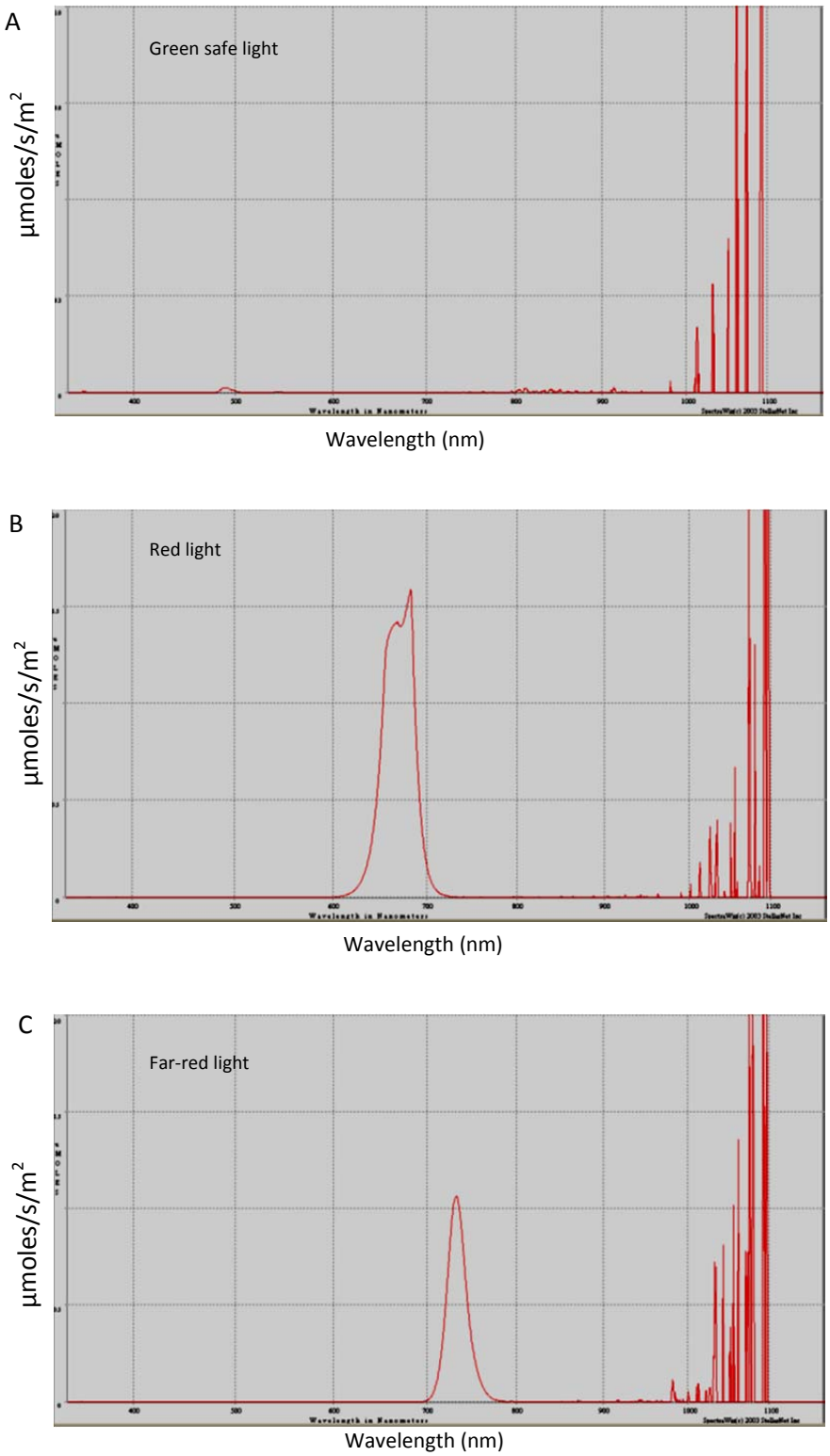


Fig 2.18: Light spectrum used during immunoprecipitation assays (A) Right under the green safe light source (B) Red light used for photoconversion of holophytochrome to Pfr (C) Far-red light used to revert Pfr form of holophytochromes back to Pr form.

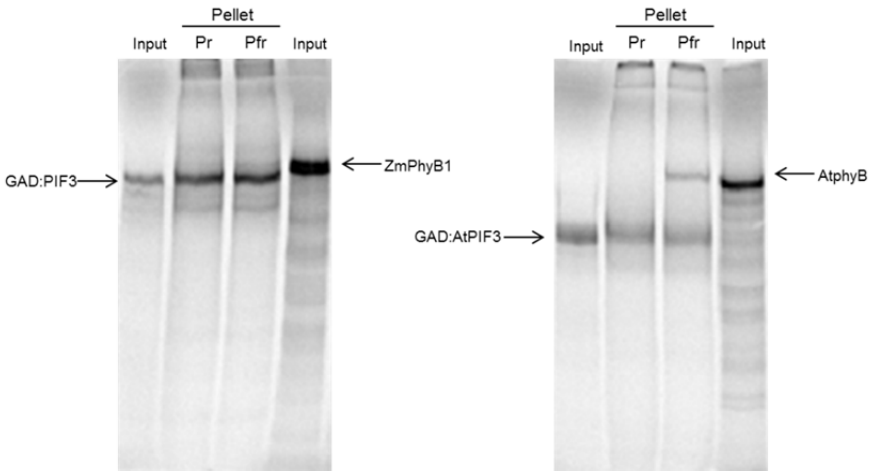


Fig 2.19: Immunoprecipitation Assays (Phosphorimaging scans of SDS-PAGE) using GAD:ZmPIF3 as bait. **Left:** No binding observed in either Pr or Pfr form of ZmphyB1. **Right:** Positive control showing interaction of AtphyB (Pfr) with AtPIF3. Pr form doesn't bind.

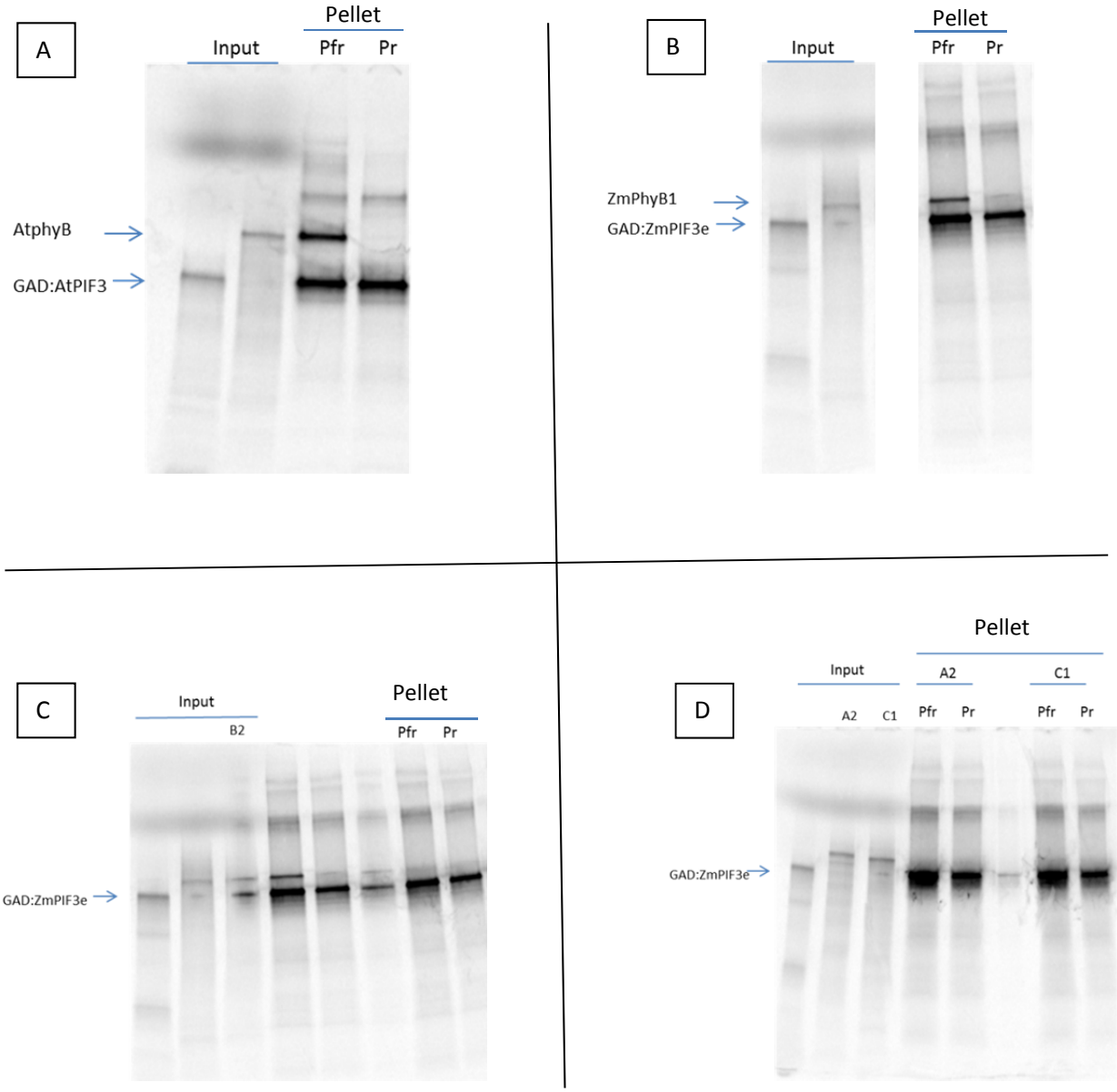


Fig 2.20: Immunoprecipitation Assays (Phosphorimaging scans of SDS-PAGE) (A) Positive control showing interaction of AtphyB (Pfr) with AtPIF3. Pr form doesn't bind. (B) Pfr form of ZmphyB1 bind to ZmPIF3e (C) Both Pr and Pfr forms of ZmphyB2 do not show any binding (D) ZmphyA2 and ZmphyC1 (Pr and Pfr) don't show binding to ZmPIF3e either.

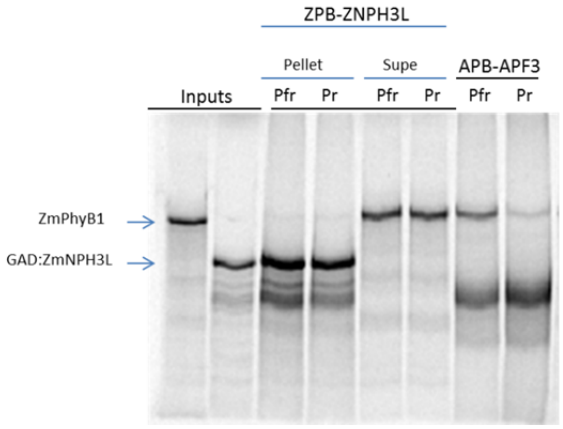


Fig 2.21: Immunoprecipitation Assay (Phosphorimaging scans of SDS-PAGE): ZmNPH3L protein did not pull down Pfr or Pr forms of ZmphyB1 (see lanes labeled as Pellet). APB-APF3: Positive control with GAD:AtPIF3 as bait and AtphyB (Pr and Pfr) as prey.

Chapter 3

The Basic Helix-Loop-Helix (bHLH) family of proteins in Maize

Abstract

bHLH (basic helix-loop-helix) proteins form one of the largest families of transcription factors in eukaryotes. Genome wide analysis of bHLH transcription factors has been done in Arabidopsis and rice but not in maize. Using bioinformatic tools, the ZmbHLH members were identified from the recently sequenced maize genome. At least 197 ZmbHLH genes were found which were characterized based on phylogenetic relationship of the bHLH domain, non-bHLH motifs, intron pattern inside bHLH domain and DNA binding properties. This study suggests functional conservation of bHLH proteins in maize compared to Arabidopsis and rice while still showing potentially novel functions of some maize bHLH proteins.

Introduction

First described in murine transcription factors E12 and E47 (Murre et al. 1989), basic helix-loop-helix (bHLH) domain proteins are found in most eukaryotes. bHLH proteins constitute one of the largest families of transcription factors in fungi, animals and plants (Riechmann et al. 2000; Ledent and Vervoort 2001). This transcription family has been shown to be involved in regulation of several key biological processes such as environment signal sensing, cell cycle control, cell proliferation, lineage establishment and epidermal differentiation in animals and responses to environmental factors, control of secondary metabolism pathways and developmental regulation in plants (Massari and Murre 2000).

The bHLHs have been extensively characterized in animals but less so in plants. Recent studies have shown involvement of plant bHLHs in diverse biological functions such as PIFs/HFR1 in light signaling (Ni et al. 1998; Fairchild et al. 2000; Huq and Quail 2002; Khanna et al. 2004; Leivar et al. 2008), OsRERJ1/AtATR2/AtICE1 in stress responses (Smolen et al. 2002; Chinnusamy et al. 2003; Kiribuchi et al. 2004), DYT1/HECATE/SPATULA in flower and fruit development (Zhang et al. 2006; Gremski et al. 2007), Arabidopsis GL3 in trichome development (Payne et al. 2000; Morohashi et al. 2007), Arabidopsis SCRM/SPCH/MUTE/FAMA in stomata development (Kanaoka et al. 2008) and LONESOME HIGHWAY(LHW) in Arabidopsis root development (Ohashi-Ito and Bergmann 2007).

The bHLH family is defined by the signature bHLH domain consisting of 60-70 amino acids. The bHLH domain is typically arranged in bifunctional regions: a basic region of 13-17 residues located at the N-terminal end and a HLH (helix-loop-helix) region at the C-terminal end of the domain. The basic region is enriched with basic amino acid residues and is involved in DNA binding. The HLH region consists of two conserved hydrophobic-residue-rich α-helices separated by a loop of variable length (Ferre-D'Amare et al. 1993); the HLH region facilitates homo/hetero -dimer formation by protein-protein interaction (Massari and Murre 2000). Each unit of the dimer has been shown to bind to half of the DNA recognition sequence primarily via the basic region (Ma et al. 1994; Shimizu et al. 1997).

bHLH proteins recognize a consensus DNA sequence motif of six nucleotides known as E-box (5'-CANNTG-3'). One common type of E-box consists of the sequence 5'-CACGTG-3'and is known as the G-box. Plant bHLH proteins have been shown to have preferential recognition of and binding to G-box elements (Martinez-Garcia et al. 2000; Huq and Quail 2002; Huq et al. 2004). The basic region of the bHLH domain contains certain conserved residues which dictate the recognition of core consensus DNA sequences, and some other residues provide specificity for the type of E-box e.g. G-box, N-box (Robinson et al. 2000).

In plants, red and far red light signaling is mediated by phytochrome photoreceptors. One primary mechanism of phytochrome mediated signaling has been found to be physical interaction of phys (Pfr form) with a subfamily of bHLH transcription factors, PIFs (PHYTOCHROME INTERACTING FACTORS) that includes PIF1, PIF3, PIF4, PIF5, PIF6, PIF7 and PIF8 (Ni et al. 1998; Ni et al. 1999; Monte et al. 2007; Shin et al. 2007; Leivar et al. 2008; Leivar

and Quail 2011). In Arabidopsis, PIF1, PIF3, PIF4, PIF5 and PIF7 have been shown to bind to G-box DNA motifs (Leivar and Quail 2011). PIFs have been implicated in various photomorphogenic functions including seed germination, promoting seedling skotomorphogenesis, regulating seedling de-etiolation, flowering and the shade avoidance syndrome (SAS) (Leivar and Quail 2011). Recently, several non-DNA binding bHLHs (atypical bHLH proteins) have been reported in Arabidopsis which lack the N-terminal basic region (Carretero-Paulet et al. 2010). Several of these atypical bHLHs have been found to negatively regulate other bHLH proteins by forming hetero-dimers. AtKDR (KIDARI) binds to AtHFR1 (long hypocotyl in far-red1) and represses light signal transduction (Hyun and Lee 2006). AtHFR1 is also an atypical bHLH which heterodimerizes with AtPIF3 to modulate phyA and cryptochrome mediated responses (Fairchild et al. 2000). PIF4 and PIF5 bHLH factors are required for shade avoidance syndrome where they bind to G-box motifs present in the promoters of shade marker genes. AtHFR1 has been shown to bind to AtPIF4 and AtPIF5 thereby forming non-DNA binding heterodimers thus inhibiting SAS (Hornitschek et al. 2009). AtPAR1 and AtPAR2 (Phytochrome Rapidly Regulated) proteins also belong to the atypical bHLH category and act as direct transcriptional repressors of two auxin-responsive genes, *SAUR15* and *SAUR68* (*SMALL AUXIN UPREGULATED*) thus negatively regulating the SAS (Roig-Villanova et al. 2007).

In animals, bHLHs have been classified into six groups, namely A-F, based on their evolutionary origin, sequence similarity, DNA binding features and function (Atchley and Fitch 1997; Ledent and Vervoort 2001). Group A bHLHs (such as MyoD, Twist) recognize and bind to the E-box motif (CACCTG/CAGCTG); Group B includes Myc, Max, SPEBP that binds to the E-box subtype CACGTG (G-box). Some group B bHLH proteins such as Myc, Mad, and SREBP also contain a

Leucine Zipper (LZ) motif, which is involved in protein dimerization. Group C (also known as bHLH-PAS) members contain an additional protein-protein interaction domain, PAS (Per/ARNT/Sim domain) and bind to non-E-box motifs such as NACGTG/NGCGTG. The PAS domain in bHLH-PAS proteins helps them to dimerize with other PAS proteins and also to interact with non-PAS proteins (Crews 1998). Group D includes bHLH members which lack a prominent basic region and are thought to be non-DNA binding proteins. These are also known as atypical bHLH proteins. Group E bHLH members bind to N-box motifs (CACGGC/CACGAC) and possess very low affinity to E-box motifs (Vandoren et al. 1994) and group F contains an additional domain involved in dimerization and DNA binding, COE (Collier/Olf-1/EBF) domain (Crozatier et al. 1996). Group F members have also been found to be quite diverged from other bHLH families (Ledent and Vervoort 2001). The group C, E and F members of the bHLH family have not been found in plant systems (Pires and Dolan 2010).

Maize is a major food crop worldwide and also one of the important model organisms in fundamental scientific research. Maize was domesticated about ten thousand years ago from the grass *teosinte* in central America (Doebley et al. 2006). Maize genome is about 2.3 Gb in size and consists of 10 chromosomes and about 85% of the genome is composed of transposable elements (Schnable et al. 2009). The maize genome has also undergone multiple rounds of genome duplication including a recent whole genome duplication event about 5 to 12 million years ago (Blanc and Wolfe 2004). Over the last century, breeders have been successful in increasing grain yields up to eightfold (Troyer 2006) by taking advantage of heterosis, the mechanism of which is not well understood. As bHLH proteins are key elements in regulation of growth and development, identification and characterization of bHLH genes in maize can

provide greater insight and can be potentially helpful in studies related to yield increase. Although the maize genome has recently been sequenced by the Maize Genome Sequencing Project (www.maizesequence.org), the maize bHLH gene family has not been analyzed yet. Interestingly, the first plant bHLH gene to be reported was maize *Lc*, a maize *R* gene family member involved in the regulation of anthocyanin biosynthesis (Ludwig et al. 1989), however a detailed analysis of the maize bHLH gene family along with the phylogenetic relationship with other plant bHLH genes remains to be performed. In this study, I identified 197 putative ZmbHLH genes across the maize genomic sequence (Maize Genome Sequence Project Release 5b). The putative ZmbHLH genes were analyzed for the phylogenetic relationship with Arabidopsis and rice bHLH subfamilies (Li et al. 2006; Pires and Dolan 2010) which was also supported by intron patterning and location within bHLH domains and estimation of DNA binding ability conferred by the basic region. Based upon the presence of an APB (Active Phytochrome Binding B) motif (motif 14), PIF family members in maize were identified. This analysis will provide a better understanding of maize transcriptional regulation and phytochrome mediated pathways.

Materials and Methods

bHLH Gene search

The maize predicted protein dataset [ZmB73_5b_FGS_translations.fasta (filtered gene set)] was obtained from www.maizesequence.org (Maize genome sequence release 5b). Putative bHLH sequences were identified by using hidden Markov model search (hmmsearch v. 3.0; http://hmmer.janelia.org/) (Eddy 1998) across maize predicted proteins dataset. The PFAM HLH hidden Markov model (PF000010_21; http://pfam.sanger.ac.uk) was used for this search using default settings. A total of 312 hits containing putative HLH domains were identified. 204 hits representing each unique locus were selected on the basis of presence of complete HLH domains or longer protein sequence. Hits with partial HLH domains were manually curated to identify mispredictions of splice sites. Missing amino acids at N-terminal or C-terminal ends of the bHLH domains were obtained from the corresponding full length protein by a Perl script (HLH_domain_extender_bothends_fixed.pl). Seven hits with incomplete HLH domain (GRMZM2G137380, GRMZM2G138454, GRMZM2G175955, GRMZM2G317317, GRMZM2G452996, GRMZM5G896413 and GRMZM2G100313) were discarded and finally 197 entries were used further analysis. Homeolog data for pairing within ZmbHLH genes was identified from a list (Maize Paralogs.xlsx) created by James Schnable of UC Berkeley (obtained from Nathan Springer group, University of Minnesota). Orthologous gene information was obtained from www.phytozome.org (Zmays_181_annotation_info.txt).

Sequence alignment and phylogenetic analysis

Multiple alignments were created using MAFFT v6 (Windows) using the –auto, --reorder options. Visualization and editing of alignments was performed using Geneious Pro v5.06 (Biomatters Ltd, Auckland, New Zealand) and MEGA v5.05. Identification of conserved residues and E-box/G-box binding motifs in the alignment was done in Geneious Pro v5.06. bHLH sequences in the alignment were renamed as ZmbHLHXXX as per (Bailey et al. 2003; Heim et al. 2003). Maximum likelihood phylogenetic analysis of the bHLH sequence alignment was performed using RAxMLHPC-MPI v 7.2.6 using the JTT model of amino acid substitutions and gamma distribution parameters estimated by the software. Regions with gaps in the alignment were removed. 1000 rapid bootstraps were performed. A neighbor joining (NJ) tree was created using MEGA v5.05 with Poisson substitution, uniform rate, pairwise deletion for gaps/missing data and 1000 bootstraps. Phylogenetic trees were visualized and edited in MEGA v5.05. For classification of ZmbHLHs into subfamilies, representative bHLH sequences for each subfamily from Arabidopsis, rice and *Physcomitrella patens* (Table 1) were added to the alignment and maximum likelihood and neighbor joining trees were created as above.

Detection of conserved motifs outside bHLH

Screening of conserved motifs outside the bHLH domain was done by hmmsearch using the hidden Markov model of each motif. Alignments of 26 conserved motifs were obtained from the supplementary material of (Pires and Dolan 2010). Hidden Markov models for motif alignments were generated by hmmbuild 2.3.2 and each model was screened across the full

length proteins of the 197 bHLH maize genes using hmmsearch. An e-value cut off of 10 was used for the search.

Intron pattern and location

Intron location and patterning inside the bHLH domain was analyzed by software tool at http://www.webscipio.org/ where intron location is determined by comparing protein sequence to the genomic sequence (Odronitz et al. 2008; Hatje et al. 2011). ZmbHLH domain sequences and the corresponding genomic sequences obtained from www.maizesequence.org (Release 5b) were used as input.

Results and Discussion

The maize genome contains at least 197 bHLH genes

By using the PFAM HLH hidden Markov model (PF000010_21), the HMMsearch tool revealed a total of 312 putative bHLH transcription factor coding genes in maize genome. The list contained multiple transcript variants from a number of loci and only one representative bHLH sequence was selected for each locus. Some mis-annotations were also found with incorrect intron-exon predictions. Genomic sequences of corresponding entries were manually analyzed for identification of correct splice sites. The small number of partial bHLH sequences was completed by examining genomic sequences downstream of the predicted coding region and identifying possible missed exons. The FGENESH gene prediction tool at www.softberry.com was utilized where required to identify missed exons. Partial bHLH sequences that could not be completed using these methods were not included in this analysis. Finally, 197 representative, complete bHLH domain sequences were aligned and further analyzed. Each sequence in the alignment was named as per convention i.e. ZmbHLHXXX (Fig 3.1) (Bailey et al. 2003; Heim et al. 2003). The list of genes indicating the corresponding bHLH numbering and other information can be found in Table 3.3. Based upon the homeologous gene list by James Schnable, a total of 100 ZmbHLH genes were found to exist as homeolog pairs (Table 3.3). The high number of homeologs was not surprising because of the recent genome duplication in maize. The bHLH signature domain consisted of approximately 60 amino acids with a typical N-terminal 13

residues long basic region rich in basic amino acids followed by two alpha-helices (approximately 14 and 21 residues in length) separated by a loop of variable length (typically 7-10 residues). The numbering and identification of regions were performed as per (Atchley and Fitch 1997; Atchley et al. 1999; Carretero-Paulet et al. 2010). Fig 3.2 shows a part of the full alignment with numbering of positions in the bHLH domain.

Key Amino acids in bHLH domain are highly conserved

The ZmbHLH sequences were analyzed for conserved residues as per the consensus motif by (Atchley et al. 1999) and also compared with the amino acid frequencies observed for the conserved locations in Arabidopsis bHLHs (Toledo-Ortiz et al. 2003). In Arabidopsis, certain positions in the bHLH region were identified that were less conserved than the consensus motif described in animals by (Atchley et al. 1999) indicating few differences between animal and plant bHLHs. I found the conserved residues at various positions in maize bHLHs to be highly similar to that in Arabidopsis, consistent with the finding of differences between animal and plant bHLHs. (Table 3.2 shows a frequency comparison of conserved residues in animal, Arabidopsis and maize bHLH domains). This suggests that the functions of bHLH proteins in Arabidopsis and maize are potentially conserved. It should be noted that there are a few differences at certain positions such as occurrence of serine (S), glutamine (Q) and histidine (H) at positions 24, 50 and 60 respectively in ZmbHLHs which are absent in AtbHLHs. It is not yet known if inclusion of these amino acids at the mentioned positions has any effect on ZmbHLH functions.

X-ray crystallography studies have shown that the basic region of the bHLH domain is primarily involved in interacting with DNA (Ferre-D'Amare et al. 1993; Atchley and Zhao 2007). Most bHLH proteins have been shown to bind 5'-CANNTG-3'; hexanucleotide sequences also known as E-boxes. E-box binding bHLH proteins possess a conserved glutamic acid (E) at position 9 of the basic region that contacts the DNA at the CA nucleotides of the E-box sequence (Ferre-D'Amare et al. 1993; Atchley et al. 1999). In plants, 74% of the analyzed bHLH proteins have been shown to possess the critical E_9 and R/K_{12} residues (Pires and Dolan 2010). Consistent with that information, 150 out of 197 (76%) maize bHLH proteins were found to have the E_9-R/K_{12} motif (Table 3.5). Amino acids at other positions in the basic region have been suggested to provide specificity for different E-box sequences (Atchley and Fitch 1997; Ledent and Vervoort 2001; Atchley and Zhao 2007). In animals, group A and group B bHLH proteins bind to specific E-box configurations CAGCTG/CACCTG and CACGTC/CATGTTC respectively and have been found to possess a characteristic arginine (R) at position 9 in group A and conserved H/K_5 and R_{13} in group B (Atchley and Zhao 2007) . In plants, 53% of bHLH proteins have been shown to have the characteristic H/K_5-E_9-R_{13} similar to group B suggesting that most are E-box binding proteins (Pires and Dolan 2010). In support of this, several plant bHLH proteins have been reported to bind to an E-box subset with CACGTG sequence also referred to as G-box motif (Martinez-Garcia et al. 2000; Toledo-Ortiz et al. 2003; Qian et al. 2008). I found that 120 of the maize bHLH sequences (~60%) have the conserved H/K_5-E_9-R_{13} motif and thus potential candidates to be G-box binding proteins (Table 3.6). N-box (CACGCG/CACGAG) binding bHLH proteins possess a proline at position 6 in addition to H_5-E_9-R_{13} and have been typically found only in animals (Atchley et al. 1999). I was also not able to find any maize bHLH sequence with this

configuration. Similar to other plants and animal bHLH sequences, maize bHLHs also contains a highly conserved R (arginine) at positions 10 and 12 with 79% and 92% respectively. A conserved Q_5-A_9-R_{13} motif has also been reported in ~11% of plant bHLH proteins and has been suggested to bind to a novel target DNA sequence instead of the E-box as it does not contain the required E residue at position 9 (Pires and Dolan 2010). Interestingly, I also found ~11% (21 out of 197) of maize bHLH sequences to have Q_5-A_9-R_{13} (Fig 3.3). One of the maize bHLH genes containing the Q_5-A_9-R_{13} motif, *ZmbHLH034* has been previously identified as *BA1* (*barren stalk 1*) shown to be required for initiation of lateral meristems (Gallavotti et al. 2004).

The α-helical region in the bHLH is involved in dimer formation. The structure of the dimer has been shown to be stabilized by conserved hydrophobic residues isoleucine (I), leucine (L) and valine (V) at conserved residues (Ferre-D'Amare et al. 1993; Atchley et al. 1999). A conserved leucine is present at positions 23 and 64 in 99% and 96% of plant bHLHs. Positions 54 and 61 contain an I, L or V in 99% and 93% of plant bHLHs (Pires and Dolan 2010). Maize bHLHs indicate similar conservation at these positions with 100% L at position 23, 99% L at 64, 88% and 94% of I/L/V at positions 54 and 61 respectively. A conserved proline (P) is also present at position 28 in maize bHLH similar to other plant and animals. This proline residue demarcates the first helix and start of the loop of variable length.

Thus, maize bHLHs also show high level of conservation of various key amino acids in the basic and helical regions suggesting a similar DNA binding, dimer forming and transcription factor activity as shown in other plants and animals.

DNA binding properties

The basic region of the bHLH domain confers the capacity for DNA binding (Massari and Murre 2000). Presence of at least 5 basic residues in the basic region has been used as a criterion to classify HLH proteins with potential to bind DNA (Massari and Murre 2000). Maize bHLH proteins can be categorized in to two main groups according the N-terminal basic region of 13 amino acids: (1) 113 ZmbHLH proteins with at least 5 basic residues present in the basic region indicating potential DNA binding capability and (2) 84 ZmbHLH proteins lacking the basic region i.e. less than 5 basic residues in the corresponding N-terminal region that are thus likely to be devoid of DNA binding properties. The DNA binding bHLH proteins can be further divided into two subgroups with specificity for different DNA motifs. Presence of E_9 (Glutamic acid) and R_{12} (Arginine) have been reported to be involved in E-box motif recognition (Ferre-D'Amare et al. 1994; Shimizu et al. 1997). A lysine (K) instead of arginine (R) at position 12 has been shown not to inhibit the E-box binding as well (Hua et al. 1993). Out of 113 predicted DNA binding ZmbHLHs, 105 (53% of total) also contain the putative E-box-binding recognition motif and the remaining 8 probably binds to non-E-box motifs. Furthermore, presence of H/K_5-E_9-R_{13} in the basic region provides specificity for binding to the G-box motif, a subtype of E-box DNA motifs (Massari and Murre 2000). A total of 79 ZmbHLHs were found to be putative G-box binding. The number of G-box binding and non-G E-box binding bHLHs in maize is similar to rice and Arabidopsis (Carretero-Paulet et al. 2010). Table 3.4 shows a classification of bHLH proteins based on its DNA binding properties and a comparison of family size in maize with Arabidopsis and rice.

bHLH proteins with a basic region containing less than 5 basic residues have been predicted to be non-DNA binding proteins. In my analysis, I found 84 ZmbHLHs to contain less than 5 basic amino acids in the basic region. However, 45 of these also contain conserved E-box binding motif out of which 41 contained the G-box binding motif (Table 3.4). It is not yet known if the less basic, E-box motif-containing bHLH proteins also have DNA binding ability. On another note, I considered 13 amino acids upstream of the first helical region to be the basic region as originally proposed by (Atchley and Fitch 1997) and used by other researchers (Heim et al. 2003; Pires and Dolan 2010) (however some other studies included 17 amino acids as the basic region (Toledo-Ortiz et al. 2003)). It is possible that inclusion of 4 additional amino acids would increase the number of basic residues thus moving some of the non-DNA binders to the DNA-binding category. Similarly, a significantly lower number of ZmbHLH was categorized as Non E-box binders (8) and a much higher number in non-DNA binders (39) compared to Arabidopsis (32 and 35) and rice (19 and 20) bHLH members (Table 3.4). Again, extending the basic region to 17 amino acids could potentially move some of the non-DNA binders to non E-box binders. 39 of the ZmbHLHs have been categorized as non-DNA binding proteins as they contain insufficient basic residues in the basic region and no E-box binding motifs as described in (Toledo-Ortiz et al. 2003). These non-DNA binding proteins are also known as “atypical” bHLH and have been shown to be involved in negative regulation of other bHLH transcription factors by protein-protein interaction in Arabidopsis (Fairchild et al. 2000; Hornitschek et al. 2009). For example, AtKDR (KIDARI) binds to and inhibits AtHFR1 function thereby regulating light mediated responses (Hyun and Lee 2006). In my analysis, I found several potential atypical bHLH proteins including ZmbHLH194 that is an ortholog to *AtKDR* (TAIR accession #

AT1G26945.1) (Table 3.3). Similarly, ZmbHLH195 is an ortholog of *AtBS1* (TAIR accession # AT1G74500.1) (Table 3.3), another atypical bHLH protein known to have a role in light signaling (Castelain et al. 2012). This suggests a likely similar non-DNA binding role by maize atypical bHLH proteins however confirmation at molecular level is required.

Phylogenetic analysis of bHLH domains

Plant bHLH proteins have been classified into multiple subfamilies based upon phylogenetic analysis using the alignment of the conserved bHLH domain sequences (Heim et al. 2003; Toledo-Ortiz et al. 2003; Li et al. 2006) and recently 26 subfamilies has been proposed (Pires and Dolan 2010). A neighbor-joining tree (Fig 3.4) and a maximum likelihood tree (Fig 3.5) for the maize bHLH region were constructed, each with 1000 bootstraps. As found in other bHLH studies, the inner nodes of the trees show lower bootstrap support values compared to the outer nodes because of the small size of the bHLH domain (Atchley and Fitch 1997). Most of the ZmbHLH members were grouped together in clades with higher bootstrap values in both ML and NJ tree analysis. Subfamily VIIa+b (PIF family members) could be grouped together using the NJ method with bootstrap value of 41 but a lower value (15) in the ML tree. Subfamilies IVa, IVb and IVc were distinguished by the NJ method, however IVb and IVc could not be separated using the ML method. Subfamily Ib(1) and Ib(2) could only be identified in the NJ tree but not in the ML tree. For identification of subfamilies, separate trees were also created while including previously classified members from each subfamily belonging to Arabidopsis, rice and *Physcomitrella patens* (Table 3.1). Most of the maize bHLHs showed up in the same branch with

the predicted Arabidopsis, rice and Physcomitrella orthologs. A cladogram of the ML tree showing the subfamily categorization based on the location of subfamily representatives in the tree is shown in Fig 3.6 thus confirming the classification of ZmbHLH members. ZmbHLH133 could not be categorized by either of the two methods and has been reported as an orphan. ZmbHLH164 appeared as an orphan in NJ but under subfamily Iva in the ML tree based classification. ZmbHLH134 and ZmbHLH135 were categorized as orphan by ML method however under subfamily Ib1 by NJ method. Overall, subfamily determination was quite similar using both methods providing high confidence. Most of the members in a subfamily shared one or more non-bHLH motifs (Table 3.3) thus supporting the subfamily classification based upon bHLH sequence, such as motif 17 and 23/24 in members of subfamily VIIIb and XII respectively. The intron pattern inside the bHLH coding region was also highly conserved within a subfamily, such as pattern I and A in subfamily VIIIb and XII respectively. Conserved intron pattern provided another support for phylogenetic subfamilies. [See further for detail on non bHLH motifs and intron pattern]

Intron/exon structure within bHLH domains

The intron distribution pattern within the bHLH domain of all maize bHLH genes was also analyzed. In plant bHLH domains, nine frequently occurring intron patterns (A through I; Fig 3.7) has been shown which consists of up to 3 introns (Toledo-Ortiz et al. 2003; Li et al. 2006; Carretero-Paulet et al. 2010). A majority of the maize bHLH domains (76%) were found to have intron pattern type A, D or B similar to rice and Arabidopsis (Toledo-Ortiz et al. 2003; Li et al.

2006). However, a significantly higher number of maize bHLH genes (11%) contain intron pattern type B compared to Arabidopsis (~5%) and rice (~4%) (Fig 3.7) (Li et al. 2006). Intron pattern type B contains the same intron positioning to the first 2 introns as type A but lacks the 3rd intron. Positioning of introns was also highly conserved as shown in Fig 3.7. The most common pattern was found to be type D with a single intron present inside the loop region. About 15% of the maize bHLH genes showed no introns within the bHLH domain (Pattern I, Fig 3.7). Previous studies have shown that the intron/exon position and distribution pattern within the bHLH domain is conserved across subfamilies (Toledo-Ortiz et al. 2003; Li et al. 2006; Carretero-Paulet et al. 2010). I analyzed individual ZmbHLH subfamilies to determine if this also holds true in maize. Indeed, most of the maize bHLHs showed conserved intron patterns within a subfamily, for example subfamily XI contains type B, subfamily VIIIb contains type I and subfamily XII contains type A intron patterns exclusively. Some exceptions were found, which included subfamily VIIa+b, VIIIc2, IX and X where more than one pattern type existed. The conserved intron patterning further supports the classification of bHLH genes into subfamilies determined by the phylogenetic analysis of bHLH region alignment. The complete list of intron patterns in bHLH domains of maize bHLH genes can be found in Table 3.3.

Conservation of non-bHLH motifs

Outside of the bHLH domain, the amino acid sequence has been found to be quite divergent in the bHLH protein sequence. However, at least 28 conserved non-bHLH motifs have been reported to be present in plant bHLH proteins (Pires and Dolan 2010). Members of the same

bHLH subfamilies have been shown to share highly conserved motifs thus supporting the subfamily classification based on bHLH domain sequences (Pires and Dolan 2010). I also identified the presence of these conserved motifs across ZmbHLH proteins. I created HMM profiles using the sequence alignments of these motifs (supplementary material (Pires and Dolan 2010)) and searched for these motifs in ZmbHLH full length protein sequences. Besides motif 15, all other motifs were found in ZmbHLHs. A complete list of maize bHLH genes with non-bHLH motif information is included in Table 3.3. Most of the maize bHLH subfamilies shared common motifs such as motifs 21, 22 and 23 in subfamily XI, motif 17 in subfamily VIIIb and motif 3 in subfamily IVc. The conservation of non bHLH motifs within a subfamily again supports the phylogenetic classification of ZmbHLH members.

Motif 14, typically present in subfamily VIIa+b members, corresponds to the active phytochrome binding (APB) motif which has been shown to be crucial for binding of PIF family members to phytochrome B in Arabidopsis (Khanna et al. 2004). There are 15 members in the VIIa+b subfamily defined by NJ and ML tree of bHLH domains in maize. My analysis showed at least 7 ZmbHLH (ZmbHLH002-008) proteins contained motif 14 / APB and are thus potential PIF family members. Previous studies have shown that subfamily VIIa+b also contain motifs 15 and 16 (Pires and Dolan 2010) however maize bHLHs completely lack motif 15 and instead contain motif 23 that is typically found in subfamily XII. ZmbHLH005 and ZmbHLH006 are the two PIF3 homeologs. Seven members of this subfamily do not contain detectable similarity to any of the described 28 motifs. In addition to the typical motif 21 and 22, subfamily XI also contains the motif 23, as seen in case of subfamily VIIa+b. Similarly, subfamily X contains motif 19 and 22 in addition to motif 20; some members of subfamily Vb contain motif 4 in addition to motif 12 and

13. In contrast, subfamily IIId+e do not contain any additional conserved detectable motifs at all. This data suggests that a domain/motif shuffling by recombination may have played a huge role in evolution of maize bHLH genes (Morgenstern and Atchley 1999). This could also have led to bHLH proteins with additional or novel functions.

I found nine bHLH proteins which contain motif 9, a leucine zipper conformation consisting of a leucine repeated at a seven-residue interval. These are also known as bHLH-ZIP proteins and have been shown to be involved in protein dimerization (Paris et al. 2003). LZ domains are typically present in members of bHLH subfamilies IVb and IVc (Pires and Dolan 2010) contiguous with the 2nd helical region of the bHLH domain (Fig 3.8). The location of LZ motif has also been reported to be same as found in animal bHLH-ZIP proteins e.g. Mad, Max, c-Myc (Atchley and Fitch 1997) however no similarities have been reported between the bHLH sequences of the plant IVb/IVc subfamily and animal bHLH-LZ proteins (Pires and Dolan 2010).

I searched for non-bHLH PFAM domains across ZmbHLH proteins. At least 6 ZmbHLH proteins (ZmbHLH85, ZmbHLH91, ZmbHLH100, ZmbHLH102, ZmbHLH130 and ZmbHLH184) were found to contain an ACT domain consistent with earlier reports in plant bHLH proteins (Pires and Dolan 2010). ACT domains are typically associated with regulatory ligand-binding proteins, mainly with metabolic roles, and have been suggested to have fused into various proteins during the course of protein evolution (Aravind and Koonin 1999; Chipman and Shaanan 2001).

Expression

Recently, a detailed whole genome transcriptome analysis of maize tissues at various developmental stages was performed using RNAseq (Davidson et al. 2011). The authors of that study identified gene expression in leaves, pre-emergence cob, post-emergence cob, mature silk, ovule, pre-emergence tassel, post-emergence tassel, whole anther, pollen, seed 5 DAP, Seed 10 DAP, embryo 25 DAP, endosperm 25 DAP. Using the FPKM data from this study (supplementary information of (Davidson et al. 2011), I checked if the predicted ZmbHLH genes are expressed. At least 163 ZmbHLH genes were found to be expressed in one or more tissues (Table 3.3) indicating a prominent role of bHLH proteins during development.

PIF (Phytochrome Interacting Factor) family

Red and far-red light induced responses in plants have been shown to be mediated by phytochromes, via interaction with the PIF family members (Leivar and Quail 2011). Phylogenetic analysis of Arabidopsis bHLHs categorized PIFs to subfamily 15 of the bHLH superfamily (Toledo-Ortiz et al. 2003). In an updated classification (Pires and Dolan 2010), PIF family members were assigned to subfamily VIIa+b that contained the characteristic APB (Active Phytochrome B binding) motif (Motif 14) (Pires and Dolan 2010). Motif 15 and 16 were also present in some of the subfamily members. In my analysis of ZmbHLHs, I found at least 8 members (ZmbHLH002-008) that contained the APB motif and are thus potential candidate members of the PIF family. Interestingly, these do not contain motif 15, however ZmbHLH003

and ZmbHLH008 showed an additional motif 23 which was absent in Arabidopsis. In a yeast two hybrid screen using ZmPHYB1-CTD as bait, the positive interacting protein turned out to be ZmbHLH005 and showed homology to AtPIF3 at least in the conserved bHLH, APB and APA motif regions and thus was referred to as ZmPIF3 (see chapter 2). As per homeologous gene information (data by James Schnable), the maize genome contains a paralog of *ZmPif3* (Maizesequence accession # GRMZM2G115960), which was identified as ZmbHLH006 in the alignment and was named as *ZmPif3-P*. Complete coding regions of these two *ZmPif3* homeologs were cloned and an interaction study with maize phytochromes confirmed their identities as PIF family members (See chapter 2). Because of very low levels of similarity outside of bHLH domain, the orthology of other maize PIF members to Arabidopsis proteins could not be determined.

Dimerization

One of the main features of bHLH proteins is dimer formation, however the residues essential for dimerization have not been characterized. The hydrophobic residues in the helical region has been suggested to regulate protein interaction properties (Massari and Murre 2000). A crystal structure of the human Max bHLH protein revealed a required leucine residue at position 23 for dimer formation (Brownlie et al. 1997). In another study, mutations at Leu-23 in helix 1 and Leu-64 in helix 2 of AtPAR1 led to complete loss or reduction of dimerization (Carretero-Paulet et al. 2010) suggesting crucial roles of these two residues. In maize, 100% and 99% of the reported bHLH sequences contain a conserved leucine at position 23 and 64

respectively which is also quite consistent with Arabidopsis bHLHs where position 23 and 64 contain 100% and 93% leucines (Toledo-Ortiz et al. 2003). This data suggests a potentially conserved dimerization function of bHLH proteins in maize. In Arabidopsis, several homo and hetero-dimer formation among the bHLH family have been reported (Toledo-Ortiz et al. 2003; Leivar and Quail 2011).

As mentioned in chapter 2, I cloned *ZmPIF3e* and *ZmPIF3-P* CDS from the maize cDNA library. In a preliminary study using co-immunoprecipitation, I investigated whether these PIF3 homeologs are capable of forming homo and hetero-dimers. GBD:PIF3 translational fusion proteins were used to precipitate the potential dimer. ZmPIF3-P showed homo-dimer formation and hetero-dimer formation with ZmPIF3e (Fig 3.9 A lanes: e-p and p-p). Bands corresponding to ZmPIF3e were indistinct in lanes e-e and p-e. This experiment needs to be repeated to produce an unambiguous result. Despite the evolutionary distance between Maize and Arabidopsis, the phytochrome-PIF system seems to be quite conserved at least in terms of phyB-PIF3 physical interaction (See chapter 2). Based on that observation, it is likely that maize and Arabidopsis PIF3 can interact with each other and form hetero-dimers. Using co-IP, I found that both ZmPIF3e and ZmPIF3-P can make hetero-dimers with AtPIF3 as shown in Fig 3.9 B (lanes e-a and p-a). It is however not known whether the Arabidopsis-maize PIF3 heterodimer would be functional in either maize or Arabidopsis signaling pathways.

Conclusion

My analysis shows that maize genome contains a large family of bHLH genes with at least 197 members and about half of these exist as homeologous pairs. This suggests that a large number of duplicated genes were retained in the maize genome after the recent whole-genome duplication event, in some cases possibly acquiring new functions or subfunctionalization. ZmbHLH domains were highly conserved and were very similar to those in Arabidopsis and rice. Conserved residues in the basic region and helical region indicated potential binding to E-box/G-box DNA motifs and protein-protein interaction. A co-immunoprecipitation study confirmed homo- and hetero- dimerization of PIF3 proteins, members of bHLH subfamily VIIa+b. ZmPIF3 proteins also hetero-dimerized with Arabidopsis PIF3 thus indicating conserved dimerization function of bHLH proteins in maize and Arabidopsis. A few ZmbHLH members do show variation at the some of the conserved amino acid sites and further studies are required to determine any effect on function. Phylogenetic analysis of bHLH domains showed clustering of the ZmbHLH members into the previously reported subfamilies along with Arabidopsis and rice bHLH subfamily representatives. This classification was also supported by conserved intron patterns inside the bHLH region and non-bHLH motifs shared by members of the same subfamily. Some of the subfamilies contain motifs typical of other subfamilies and thus indicate a domain/motif shuffling during evolution. This process may have led to acquired novel functions by duplicated ZmbHLH genes. Most of the ZmbHLH genes were found to be expressed in one or more tissues however biological functions of only a few ZmbHLH proteins have been

characterized so far. The bHLH family is one of the largest transcription factor families involved in plant development and this study would provide foundation for further research potentially useful for crop improvement.

Figures and Tables

Ia	OsbHLH044, AtbHLH096, PpbHLH018
Ib(1)	OsbHLH144, AtbHLH095, PpbHLH037
Ib(2)	OsbHLH147, AtbHLH126, AtbHLH100
II	OsbHLH141, AtbHLH010, PpbHLH092
IIIa+c	OsbHLH006, AtbHLH029, PpbHLH041
IIIb	OsbHLH001, AtbHLH061, PpbHLH021
IIId+e	OsbHLH008, AtbHLH003, PpbHLH002
IIIf	OsbHLH012, AtbHLH012, PpbHLH004
Iva	OsbHLH018, AtbHLH025, PpbHLH022
Ivb	OsbHLH061, AtbHLH121, PpbHLH039
Ivc	OsbHLH057, AtbHLH115, PpbHLH014
Ivd	OsbHLH024, AtbHLH041, PpbHLH005
IX	OsbHLH109, AtbHLH128, PpbHLH006
Va	OsbHLH031, AtbHLH141, PpbHLH046
Vb	OsbHLH035, AtbHLH030,
VIIa+b	OsbHLH101, AtbHLH008, AtbHLH026, PpbHLH009
VIIIa	OsbHLH178, AtbHLH117, PpbHLH074
VIIIb	OsbHLH117, AtbHLH140, PpbHLH026
VIIIc(1)	OsbHLH125, AtbHLH086, PpbHLH033
VIIIc(2)	OsbHLH128, AtbHLH054, AtbHLH084, PpbHLH028
X	OsbHLH065, AtbHLH110,
XI	OsbHLH096, AtbHLH066, PpbHLH012
XII	OsbHLH079, AtbHLH078, PpbHLH001
XIII	OsbHLH149, AtbHLH156,
XIV	OsbHLH138, AtbHLH142,
XV	OsbHLH153, AtbHLH135,

Table 3.1: Representative bHLH members from each phylogenetic subfamily in Arabidopsis, rice and *Physcomitrella* that were used in multiple sequence alignment along with ZmbHLH sequences for subfamily classification in maize (Pires and Dolan 2010).

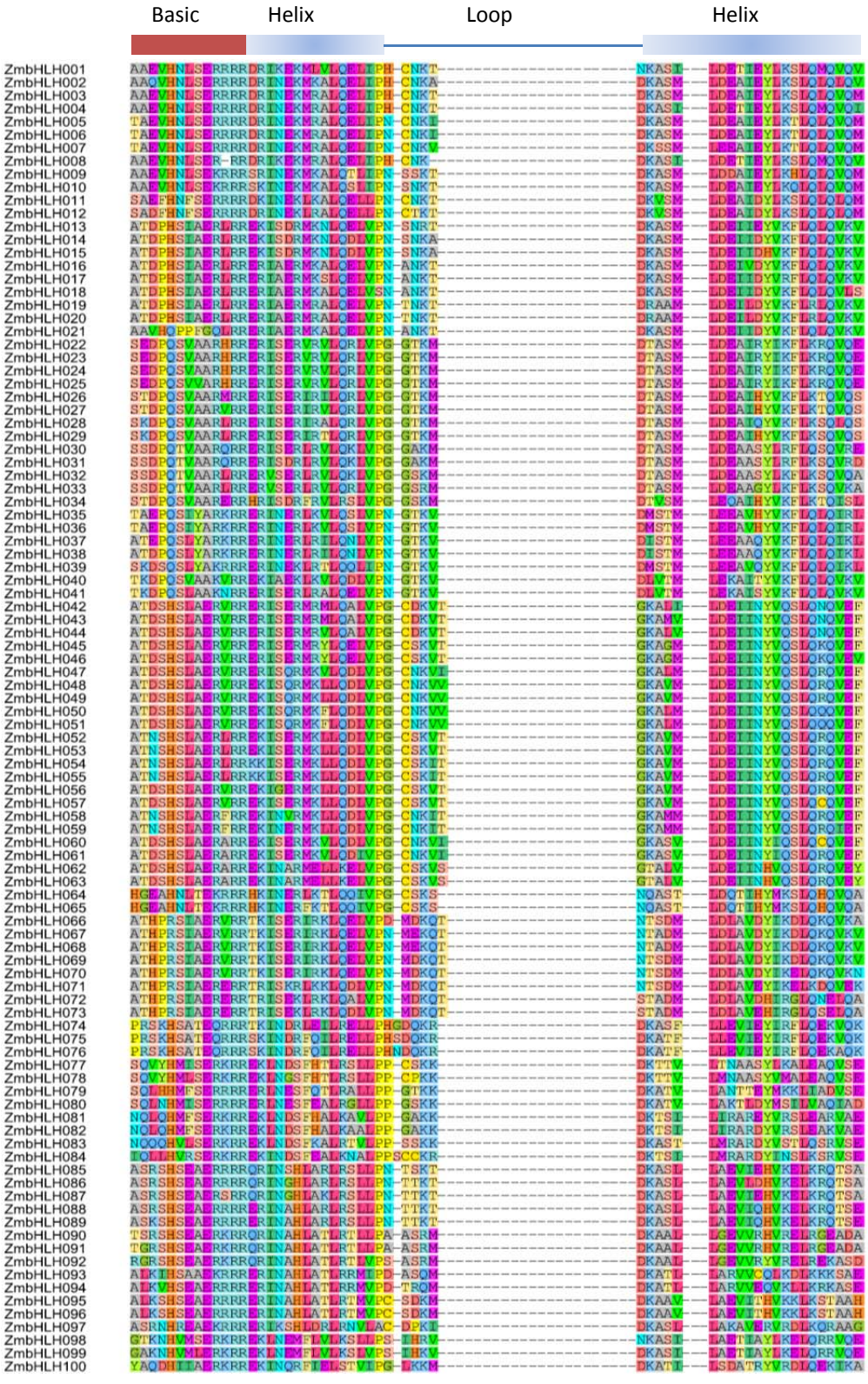
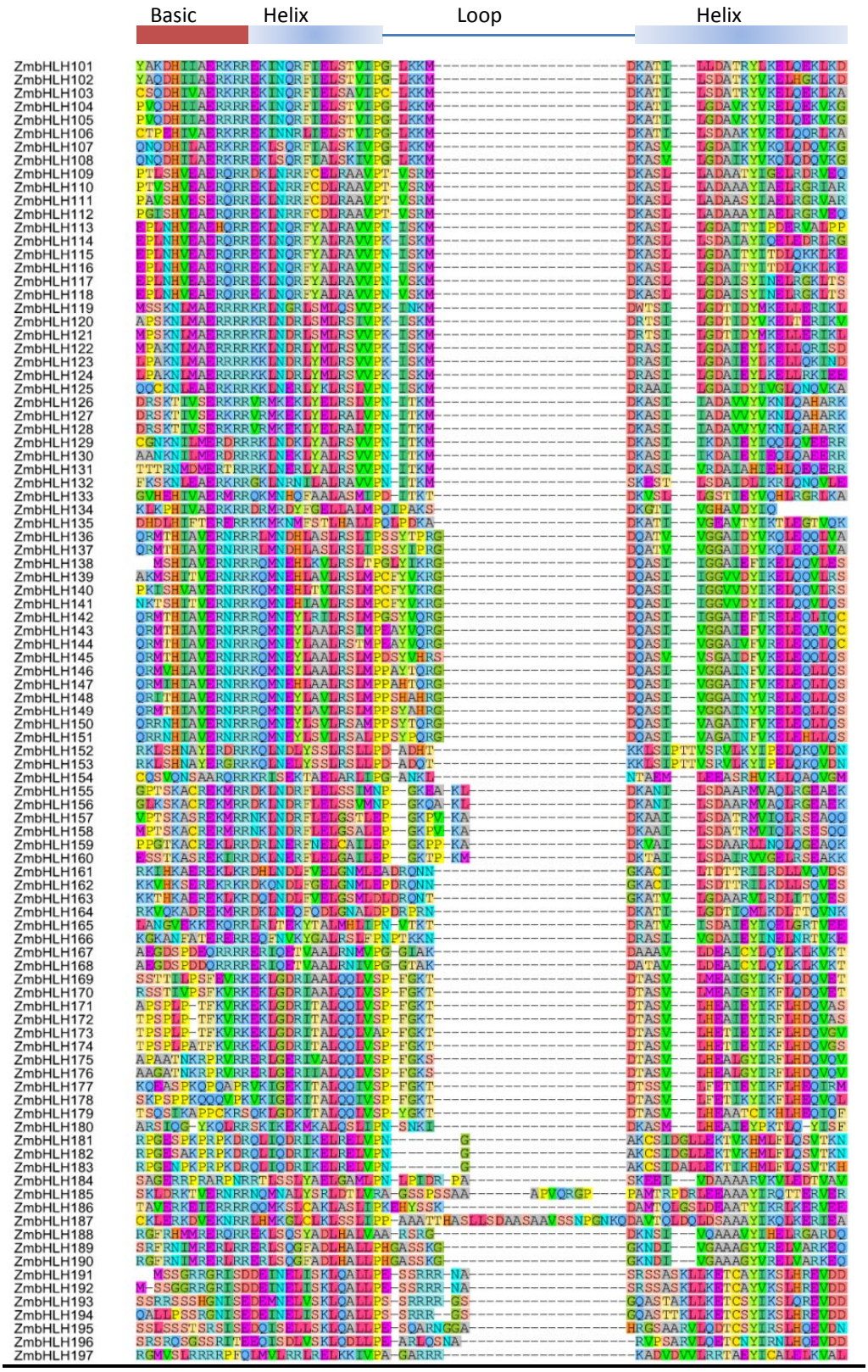


Fig 3.1: Alignment of bHLH domains from 197 putative bHLH genes in maize showing the conserved basic and helical regions with a variable loop.

Fig 3.1 (cont.)


Basic
Helix
Loop
Helix
ZmbHLH101
ZmbHLH102
ZmbHLH103
ZmbHLH104
ZmbHLH105
ZmbHLH106
ZmbHLH107
ZmbHLH108
ZmbHLH109
ZmbHLH110
ZmbHLH111
ZmbHLH112
ZmbHLH113
ZmbHLH114
ZmbHLH115
ZmbHLH116
ZmbHLH117
ZmbHLH118
ZmbHLH119
ZmbHLH120
ZmbHLH121
ZmbHLH122
ZmbHLH123
ZmbHLH124
ZmbHLH125
ZmbHLH126
ZmbHLH127
ZmbHLH128
ZmbHLH129
ZmbHLH130
ZmbHLH131
ZmbHLH132
ZmbHLH133
ZmbHLH134
ZmbHLH135
ZmbHLH136
ZmbHLH137
ZmbHLH138
ZmbHLH139
ZmbHLH140
ZmbHLH141
ZmbHLH142
ZmbHLH143
ZmbHLH144
ZmbHLH145
ZmbHLH146
ZmbHLH147
ZmbHLH148
ZmbHLH149
ZmbHLH150
ZmbHLH151
ZmbHLH152
ZmbHLH153
ZmbHLH154
ZmbHLH155
ZmbHLH156
ZmbHLH157
ZmbHLH158
ZmbHLH159
ZmbHLH160
ZmbHLH161
ZmbHLH162
ZmbHLH163
ZmbHLH164
ZmbHLH165
ZmbHLH166
ZmbHLH167
ZmbHLH168
ZmbHLH169
ZmbHLH170
ZmbHLH171
ZmbHLH172
ZmbHLH173
ZmbHLH174
ZmbHLH175
ZmbHLH176
ZmbHLH177
ZmbHLH178
ZmbHLH179
ZmbHLH180
ZmbHLH181
ZmbHLH182
ZmbHLH183
ZmbHLH184
ZmbHLH185
ZmbHLH186
ZmbHLH187
ZmbHLH188
ZmbHLH189
ZmbHLH190
ZmbHLH191
ZmbHLH192
ZmbHLH193
ZmbHLH194
ZmbHLH195
ZmbHLH196
ZmbHLH197

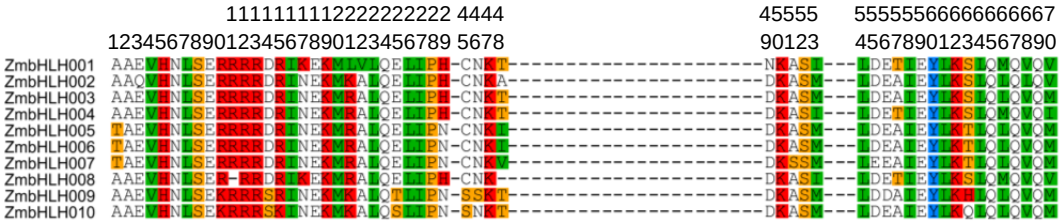


Fig 3.2: A region from the full alignment of maize bHLH sequences showing the numbering of amino acid positions as described by (Atchley and Fitch 1997; Pires and Dolan 2010).

Position	Consensus motif amino acid frequency within bHLH domains in animals (Atchley et al. 1999)	Amino acid frequency in AtbHLH domains (Toledo-Ortiz et al. 2003)	Amino acid frequency in ZmbHLH
9	E:93%	E: 76%, **A:10%**	E:76% ,**A:11%**
10	R:81%, K:14%	R:74%, K:14%	R:79%, K:13%
12	R:91%	R:91%	R:92%
16	I:35%, L:33%, V:23%	I:52%, L:27%, **M:12%**, V:3%	I:55%, L:29%, **M:13%,** V:1.5%
17	N:74%	N:51%, **S:19%**	N:55%, **S:26%**
20	F:72%	F:26%, I:14%, L:26%, **M:20%**	F:26%, I:14%, L:32%, **M:21%**
23	L:98%	L:100%	L:100%
24	K:35%, R:44%	K:4%, R:35%, **Q:42%,** G:4%	K:4%, R:35%, **Q:45%, S:6%,** G:4%
50	K:93%	K:45%, **T:13%**	K:58%, **T:18%, Q:10%,**
53	I:74%, T:15%, V:7%	I:27%, T:4%, V:16%, **L:14%**, **M:33%**	I:32%, T:3%, V:17%, **L:11%, M:32%**
54	L:98%	L:76%, **V:14%**	L:84%, **V:11%**
57	A:76%	A:60%, **I:16%**, **V:12%,** T:9%	A:59%, **I:16%, V:10%,** T:15%
60	Y:77%	Y:78%	Y:78%, **H:9.6%**
61	I:69%, L:16%, V:8%	I:40%, L:13%, V:33%	I:33%, L:16%, V:46%
64	L:80%, M:7%	L:93%, M:1%	L:99%

Table 3.2: Amino acid conservation at different positions in bHLH domain: comparison between the original bHLH consensus motif (animals), Arabidopsis and maize. Boldface letters indicate the residues that are different between animal and plant bHLH members. In comparison to Arabidopsis, few maize bHLH members also possess serine (S), glutamine (Q) and histidine (H) at positions 24, 50 and 60 respectively.

bHLH No.	MaizeGDB #	Gene Name	Homeolog	Chr No.	Start (Mb)	End (Mb)	Subfamily	Motifs	Intron Pattern	DNA Binding	Expression	At Orthlog	Os Ortholog
ZmbHLH001	GRMZM5G899865			1	267.563474	267.565495	VIIa+b (?)		B	BG	Y	AT1G75580.1	LOC_Os06g04590.1
ZmbHLH002	GRMZM5G865967			1	172.878324	172.880857	VIIa+b (?)	14	B	BG	Y	AT2G43010.1	LOC_Os12g41650.2
ZmbHLH003	GRMZM2G165042			1	252.374241	252.377817	VIIa+b (?)	14, 16	A	BG	Y	AT2G20180.2	LOC_Os03g43810.1
ZmbHLH004	GRMZM2G065374		GRMZM2G016756	1	283.599877	283.604922	VIIa+b (?)	14, 16, 23	A	BG	Y	AT3G59060.2	LOC_Os03g56950.1
ZmbHLH005	GRMZM2G387528	*Pif3*	GRMZM2G115960	8	0.935667	0.938714	VIIa+b (?)	14	A	BG	Y	AT1G09530.1	LOC_Os05g04740.2
ZmbHLH006	GRMZM2G115960	*Pif3* paralog	GRMZM2G387528	3	48.804985	48.808667	VIIa+b (?)	14	A	BG	Y	AT1G09530.1	LOC_Os05g04740.1
ZmbHLH007	GRMZM2G062541			10	87.690145	87.692179	VIIa+b (?)	14	D	BG	Y	AT1G09530.1	LOC_Os05g04740.2
ZmbHLH008	GRMZM2G016756		GRMZM2G065374	5	5.1016	5.105442	VIIa+b (?)	14, 16, 23	B	E non G	Y	AT2G43010.1	LOC_Os03g56950.1
ZmbHLH009	GRMZM2G030744			5	214.427886	214.432103	VIIa+b (?)		B	BG	Y	AT4G36930.1	LOC_Os02g56140.1
ZmbHLH010	GRMZM2G017349			9	16.485686	16.489619	VIIa+b (?)		A	BG	Y	AT4G36930.1	LOC_Os06g06900.1
ZmbHLH011	GRMZM2G080054	*spatula*	GRMZM2G042920	2	9.445775	9.44765	VIIa+b (?)		A	BG	Y	AT2G20180.1	LOC_Os04g52770.1
ZmbHLH012	GRMZM2G042920		GRMZM2G080054	10	142.714714	142.71674	VIIa+b (?)		A	BG	Y	AT4G36930.1	LOC_Os04g52770.1
ZmbHLH013	GRMZM5G839518			5	174.934497	174.938224	XI		B	AG	Y		
ZmbHLH014	GRMZM2G116785		GRMZM2G024530	6	107.382342	107.38591	XI		B	AG	Y		LOC_Os06g09370.1
ZmbHLH015	GRMZM2G024530	*ptf1*	GRMZM2G116785	9	11.560217	11.565193	XI	21, 22, 23	B	AG	Y	AT5G58010.1	LOC_Os06g09370.1
ZmbHLH016	GRMZM2G350165		GRMZM2G027563	2	180.661038	180.664197	XI	21, 22, 23	B	AG	Y	AT2G24260.1	LOC_Os09g25040.1
ZmbHLH017	GRMZM2G067654			9	12.981646	12.983821	XI	21, 22	B	AG		AT2G24260.1	LOC_Os02g55250.1
ZmbHLH018	GRMZM2G027563		GRMZM2G350165	7	115.274895	115.277443	XI	21, 22, 23	B	AG	Y	AT2G24260.1	LOC_Os09g25040.1
ZmbHLH019	GRMZM2G162382			7	10.694109	10.698145	XI	21, 22, 23	B	AG	Y	AT4G02590.1	LOC_Os03g58330.1
ZmbHLH020	GRMZM2G004356			5	4.217351	4.222264	XI	21, 22, 23	B	AG	Y	AT4G02590.1	LOC_Os03g58330.1
ZmbHLH021	GRMZM2G316758		GRMZM5G832135	4	174.002078	174.004537	XI		B	Non binder	Y	AT5G58010.1	LOC_Os02g55250.1
ZmbHLH022	GRMZM5G818776			4	39.724026	39.724889	VIIIb	17	I	Non binder	Y	AT4G00120.1	LOC_Os08g01700.1
ZmbHLH023	GRMZM5G817854			3	159.485166	159.485849	VIIIb	17	I	Non binder	Y	AT5G09750.1	LOC_Os08g01700.1
ZmbHLH024	GRMZM5G812883			4	39.742776	39.743642	VIIIb	17	I	Non binder	Y	AT4G00120.1	LOC_Os08g01700.1
ZmbHLH025	GRMZM2G354618			4	39.707389	39.708486	VIIIb	17	I	Non binder	Y	AT4G00120.1	LOC_Os08g01700.1
ZmbHLH026	GRMZM2G471635		GRMZM2G134735	4	57.34866	57.349385	VIIIb	17	I	Non binder		AT3G50330.1	LOC_Os08g36740.1
ZmbHLH027	GRMZM2G134735		GRMZM2G471635	1	199.070433	199.071201	VIIIb	17	I	Non binder	Y	AT5G67060.1	LOC_Os08g36740.1
ZmbHLH028	GRMZM2G414252		GRMZM2G045431	2	184.966846	184.968211	VIIIb	17	I	Non binder	Y	AT5G67060.1	LOC_Os09g28210.1
ZmbHLH029	GRMZM2G045431		GRMZM2G414252	7	125.32026	125.32229	VIIIb	17	I	Non binder	Y	AT5G67060.1	LOC_Os09g28210.1
ZmbHLH030	GRMZM2G164341		GRMZM2G081816	8	138.262239	138.264389	VIIIb	17	I	Non binder	Y	AT3G21330.1	LOC_Os01g38610.1
ZmbHLH031	GRMZM2G081816		GRMZM2G164341	3	219.181998	219.184295	VIIIb	16, 17	I	Non binder	Y	AT3G21330.1	LOC_Os01g38610.1
ZmbHLH032	GRMZM2G111666			3	204.462023	204.464579	VIIIb	17	I	Non binder	Y	AT3G21330.1	LOC_Os01g51140.2
ZmbHLH033	GRMZM2G094892			6	1618633	1618650 47	VIIIb	17	I	Non binder	Y	AT3G21330.1	LOC_Os05g46370.1
ZmbHLH034	GRMZM2G397518	*ba1* (*barren stalk 1*)		3	183.084299	183.085268	VIIIb		I	Non binder	Y	AT5G01310.1	LOC_Os01g61480.1
ZmbHLH035	GRMZM2G395549			3	113.433909	113.47138	VIIIc2		E	Non binder	Y	AT3G11130.1	LOC_Os11g01380.1
ZmbHLH036	AC198518.3_FG005			3	107.371836	107.373182	VIIIc2		E	Non binder		AT2G14760.1	LOC_Os03g42100.1
ZmbHLH037	GRMZM2G383841		GRMZM2G057260	2	208.616426	208.617726	VIIIc2	19	A	Non binder		AT4G33880.1	LOC_Os07g39940.1
ZmbHLH038	GRMZM2G057260		GRMZM2G383841	7	162.07966	162.085372	VIIIc2	19	A	Non binder	Y	AT2G33340.1	LOC_Os07g39940.1
ZmbHLH039	GRMZM2G165188			4	4.748404	4.750019	VIIIc2		A	Non E		AT2G14760.1	LOC_Os11g41640.1
ZmbHLH040	GRMZM2G066057			8	8.41608	8.417957	VIIIcl	18	D	Non binder		AT5G37800.1	LOC_Os01g02110.1
ZmbHLH041	AC216731.3_FG001			5	201.866686	201.867888	VIIIcl			Non binder		AT1G66470.1	LOC_Os02g48060.1
ZmbHLH042	GRMZM5G882527		GRMZM2G145579	1	195.720311	195.722763	XII		A	AG	Y		LOC_Os08g42470.1
ZmbHLH043	GRMZM2G145579		GRMZM5G882527	4	47.591068	47.593467	XII	23	A	AG	Y	AT5G50915.1	LOC_Os08g42470.1
ZmbHLH044	GRMZM2G178182	*gbof-1*		2	192.783267	192.785778	XII	23	A	AG	Y	AT5G50915.1	LOC_Os09g33580.1
ZmbHLH045	GRMZM2G176289		GRMZM2G101350	4	159.993146	159.995975	XII	19, 23	E	AG	Y	AT4G34530.1	LOC_Os02g47660.1
ZmbHLH046	GRMZM2G101350		GRMZM2G176289	5	200.937343	200.940145	XII	23		AG	Y	AT4G34530.1	LOC_Os02g47660.1
ZmbHLH047	GRMZM5G828396			1	192.03223	192.03493	XII	23	A	AG	Y	AT5G48560.1	LOC_Os08g41320.1
ZmbHLH048	GRMZM2G180406		GRMZM2G030762	3	162.063107	162.066386	XII	23, 24	A	AG	Y	AT5G48560.1	LOC_Os01g68700.2
ZmbHLH049	GRMZM2G030762		GRMZM2G180406	8	162.55742	162.560576	XII	24	A	AG	Y	AT1G68920.1	LOC_Os01g68700.2

Table 3.3: List of putative ZmbHLH genes showing corresponding gene id and known gene names, chromosome location information, phylogenetic subfamily, motif, intron pattern, predicted DNA binding feature, expression status (Y: yes) and Arabidopsis and rice ortholog. Phylogenetic subfamily classification (based upon bHLH domain) is also supported by conserved motif, intron pattern and DNA binding feature.

Table 3.3 (cont.)

bHLH No.	MaizeGDB #	Gene Name	Homeolog	Chr No.	Start (Mb)	End (Mb)	Subfamily	Motifs	Intron Pattern	DNA Binding	Expression	At Orthlog	Os Ortholog
ZmbHLH050	GRMZM2G144275		GRMZM2G083504	7	137.485911	137.488421	XII	24	A	AG	Y	AT1G68920.1	LOC_Os09g32510.5
ZmbHLH051	GRMZM2G083504		GRMZM2G144275	2	191.378248	191.380925	XII	23, 24	A	AG	Y	AT3G07340.1	LOC_Os09g32510.1
ZmbHLH052	GRMZM5G818643		GRMZM2G137541	7	129.962802	129.966073	XII	23, 24	A	AG	Y	AT1G68920.3	LOC_Os09g29830.2
ZmbHLH053	GRMZM2G137541		GRMZM5G818643	2	188.403767	188.407063	XII	23, 24	A	AG	Y	AT1G68920.1	LOC_Os09g29830.2
ZmbHLH054	GRMZM2G385543		GRMZM2G092091	5	9.118885	9.120925	XII	23, 24	A	AG	Y	AT1G68920.3	LOC_Os03g51910.1
ZmbHLH055	GRMZM2G092091		GRMZM2G385543	1	272.043212	272.047015	XII	23, 24	A	AG	Y	AT1G10120.1	LOC_Os03g51910.1
ZmbHLH056	GRMZM2G080168			4	17.234489	17.237004	XII	23	A	AG	Y	AT3G07340.1	LOC_Os11g25560.1
ZmbHLH057	GRMZM2G171818			2	220.669876	220.672227	XII	23	A	AG	Y	AT3G07340.1	LOC_Os11g25560.1
ZmbHLH058	GRMZM2G407119		GRMZM2G180452	9	144.807036	144.810304	XII	23	A	AG	Y	AT3G07340.1	LOC_Os03g12940.1
ZmbHLH059	GRMZM2G180452		GRMZM2G407119	1	29.259499	29.263235	XII	23	A	AG	Y	AT1G10120.1	LOC_Os03g12940.2
ZmbHLH060	GRMZM2G137358			1	289.344713	289.347971	XII	23	A	AG	Y	AT1G59640.1	LOC_Os03g58830.1
ZmbHLH061	GRMZM2G008898			2	232.380911	232.384091	XII	20, 23	A	AG	Y	AT1G59640.1	LOC_Os07g09590.1
ZmbHLH062	GRMZM2G378653		GRMZM2G074438	9	101.197226	101.203564	XII	23	A	AG	Y	AT2G42300.1	LOC_Os06g41060.1
ZmbHLH063	GRMZM2G074438		GRMZM2G378653	6	98.622638	98.628187	XII	23	A	AG	Y	AT2G42300.1	LOC_Os06g41060.1
ZmbHLH064	GRMZM2G165090			1	168.42818	168.432944	VIIa+b (?)		B	BG		AT1G09530.1	LOC_Os12g40710.1
ZmbHLH065	GRMZM2G018472			1	168.411591	168.416268	VIIa+b (?)		B	BG		AT2G20180.1	LOC_Os12g40730.1
ZmbHLH066	GRMZM5G879527			1	188.03161	188.0346	IX		A	E non G	Y	AT2G42280.1	LOC_Os08g39630.1
ZmbHLH067	GRMZM2G417597		GRMZM2G085467	2	189.701055	189.703654	IX		B	E non G	Y	AT2G42280.1	LOC_Os08g39630.1
ZmbHLH068	GRMZM2G085467		GRMZM2G417597	7	132.889939	132.892494	IX		B	E non G	Y	AT2G42280.1	LOC_Os09g31300.1
ZmbHLH069	GRMZM2G153454		GRMZM2G142932	7	156.738411	156.741186	IX		B	E non G	Y	AT2G42280.1	LOC_Os08g39630.1
ZmbHLH070	GRMZM2G142932		GRMZM2G153454	2	204.767728	204.771974	IX		A	E non G	Y	AT2G42280.1	LOC_Os08g39630.1
ZmbHLH071	GRMZM2G114873			3	166.448732	166.452009	IX	7	A	E non G	Y	AT1G05805.1	LOC_Os01g67480.1
ZmbHLH072	GRMZM2G437481			10	126.652215	126.656003	IX	7	A	E non G		AT1G05805.1	LOC_Os04g41229.1
ZmbHLH073	GRMZM2G137426			5	182.085577	182.09227	IX	7	A	E non G	Y	AT1G05805.1	LOC_Os02g39140.1
ZmbHLH074	GRMZM2G317450			5	204.517824	204.522813	Va	10	G	BG	Y	AT5G08130.1	LOC_Os02g49480.1
ZmbHLH075	GRMZM2G089501		GRMZM2G009478	7	130.237175	130.244865	Va	10, 11	G	BG	Y	AT5G08130.2	LOC_Os09g29930.1
ZmbHLH076	GRMZM2G009478		GRMZM2G089501	2	188.529569	188.535342	Va	10, 11	G	BG	Y	AT1G69010.1	LOC_Os09g29930.1
ZmbHLH077	GRMZM5G849600		GRMZM2G463133	3	219.828374	219.833646	Ivd		D	BG	Y	AT5G56960.1	LOC_Os01g39330.1
ZmbHLH078	GRMZM2G463133		GRMZM5G849600	8	138.938542	138.943955	Ivd		D	BG		AT5G56960.1	LOC_Os01g39330.1
ZmbHLH079	GRMZM2G340177			9	79.118794	79.123438	Ivd		D	BG		AT5G56960.1	LOC_Os06g37410.1
ZmbHLH080	GRMZM2G170559			5	155.094725	155.097958	Ivd		D	BG		AT5G56960.1	LOC_Os02g12820.1
ZmbHLH081	GRMZM2G397755		GRMZM2G132550	3	5.776179	5.778416	Ivd		D	BG		AT5G56960.1	LOC_Os01g09990.2
ZmbHLH082	GRMZM2G333582			8	2.1907478	2.1910933	Ivd		D	BG		AT5G56960.1	LOC_Os01g09990.1
ZmbHLH083	GRMZM2G175480			8	92.199625	92.200983	Ivd		D	BG		AT5G56960.1	LOC_Os01g09900.1
ZmbHLH084	GRMZM2G132550		GRMZM2G397755	6	132.187682	132.190029	Ivd		D	BG		AT5G56960.1	LOC_Os01g09990.1
ZmbHLH085	GRMZM2G526668		GRMZM2G107276	4	82.641312	82.643946	Vb	12, 13	D	BG	Y	AT1G68810.1	LOC_Os08g33590.1
ZmbHLH086	GRMZM2G176141			2	180.155096	180.158005	Vb	12, 13	D	BG	Y	AT1G68810.1	LOC_Os09g24490.1
ZmbHLH087	GRMZM2G107276		GRMZM2G526668	1	214.152598	214.154902	Vb		D	BG	Y	AT1G68810.1	LOC_Os08g33590.1
ZmbHLH088	GRMZM2G072376		GRMZM2G043854	1	36.916664	36.917486	Vb		D	BG	Y	AT1G68810.1	LOC_Os03g15440.1
ZmbHLH089	GRMZM2G043854		GRMZM2G072342	9	141.434879	141.438278	Vb	12, 13	D	BG	Y	AT3G25710.1	LOC_Os03g15440.1
ZmbHLH090	GRMZM2G145146		GRMZM2G113257	7	127.585573	127.586752	Vb	4, 12, 13	D	BG	Y	AT3G56770.1	LOC_Os09g28900.1
ZmbHLH091	GRMZM2G113257		GRMZM2G082630	2	187.124455	187.12554	Vb	4, 12, 13	D	BG	Y	AT3G25710.1	LOC_Os09g28900.1
ZmbHLH092	GRMZM2G082630		GRMZM2G113257	1	200.412732	200.413581	Vb	4, 12, 13	D	BG	Y	AT3G56770.1	LOC_Os08g37290.1
ZmbHLH093	GRMZM2G369629			6	124.197463	124.198965	Vb	12, 13	D	BG	Y	AT2G40200.1	LOC_Os05g07120.2
ZmbHLH094	GRMZM2G088443			3	14.928544	14.930691	Vb	12, 13	D	BG	Y	AT2G40200.1	LOC_Os01g06640.1
ZmbHLH095	GRMZM2G166642			3	3.342184	3.344508	Vb	12, 13	D	BG	Y	AT2G40200.1	LOC_Os01g11910.1
ZmbHLH096	GRMZM2G016039			4	63.752535	63.754892	Vb	12, 13	D	BG	Y	AT2G40200.1	LOC_Os01g11910.1
ZmbHLH097	GRMZM2G027068			1	291.901463	291.903534	Vb	12, 13, 17	D	BG	Y	AT2G41130.1	LOC_Os03g59670.1
ZmbHLH098	GRMZM5G822829	*r1(colored 1)*	GRMZM2G172795	10	138.462252	138.471072	IIIf	8	D	BG	Y	AT1G63650.1	LOC_Os04g47080.1

Table 3.3 (cont.)

bHLH No.	MaizeGDB #	Gene Name	Homeolog	Chr No.	Start (Mb)	End (Mb)	Subfamily	Motifs	Intron Pattern	DNA Binding	Expression	At Orthlog	Os Ortholog
ZmbHLH099	GRMZM2G172795	*b1*	GRMZM5G822829	2	18.41737	18.421827	IIIf	8	D	BG	Y	AT5G41315.1	LOC_Os04g47080.1
ZmbHLH100	GRMZM2G355469			10	118.369285	118.370757	Iva		D	BG	Y	AT4G37850.1	LOC_Os03g46860.1
ZmbHLH101	GRMZM2G155043			1	257.785005	257.792471	Iva		D	BG		AT4G37850.1	LOC_Os03g46860.1
ZmbHLH102	GRMZM2G120021			1	257.732056	257.737608	Iva		D	BG		AT4G37850.1	LOC_Os03g46860.1
ZmbHLH103	GRMZM2G095899			1	257.575165	257.578046	Iva		D	BG		AT4G37850.1	LOC_Os03g46860.1
ZmbHLH104	GRMZM2G068604			1	178.428791	178.433621	Iva		D	BG	Y	AT4G37850.1	LOC_Os12g43620.1
ZmbHLH105	GRMZM2G049686			1	178.438556	178.444036	Iva		D	BG	Y	AT4G37850.1	LOC_Os12g43620.1
ZmbHLH106	GRMZM2G036554			10	87.419521	87.421049	Iva		D	BG	Y	AT2G22750.2	LOC_Os03g46860.1
ZmbHLH107	GRMZM2G082343		GRMZM2G019806	5	9.616539	9.619038	Iva		D	BG	Y	AT4G37850.1	LOC_Os03g51580.1
ZmbHLH108	GRMZM2G019806		GRMZM2G082343	1	270.980401	270.982751	Iva		D	BG	Y	AT4G37850.1	LOC_Os03g51580.1
ZmbHLH109	GRMZM2G472671		GRMZM2G038479	1	196.806199	196.807431	IIId+e		I	AG	Y	AT4G17880.1	LOC_Os09g34330.1
ZmbHLH110	GRMZM2G107560		GRMZM2G013688	2	193.882051	193.882939	IIId+e		I	AG		AT1G32640.1	LOC_Os09g34330.1
ZmbHLH111	GRMZM2G013688		GRMZM2G107560	7	139.724235	139.725164	IIId+e		I	AG	Y	AT1G32640.1	LOC_Os09g34330.1
ZmbHLH112	GRMZM2G038479		GRMZM2G472671	4	194.911377	194.912204	IIId+e		I	AG		AT5G46760.1	LOC_Os09g34330.1
ZmbHLH113	GRMZM2G303463			2	227.2645	227.26856	IIId+e	7	I	AG	Y	AT1G01260.1	LOC_Os01g13460.1
ZmbHLH114	GRMZM2G159937			3	204.643909	204.645969	IIId+e	7	I	AG	Y	AT1G01260.1	LOC_Os01g50940.1
ZmbHLH115	GRMZM2G148723		GRMZM2G076636	3	1.600739	1.613176	IIId+e		I	AG	Y	AT5G11390.1	LOC_Os05g15140.1
ZmbHLH116	GRMZM2G076636		GRMZM2G148723	8	29.230806	29.234094	IIId+e	7	I	AG	Y	AT1G01260.1	LOC_Os01g13460.1
ZmbHLH117	GRMZM2G049229		GRMZM2G001930	9	112.004919	112.007822	IIId+e	7	I	AG	Y	AT4G17880.1	LOC_Os10g42430.1
ZmbHLH118	GRMZM2G001930	*myc7e*	GRMZM2G049229	1	98.354261	98.357077	IIId+e	7	I	AG	Y	AT1G32640.1	LOC_Os10g42430.1
ZmbHLH119	GRMZM2G435015			7	140.22562	140.228624	IIIb		D	E non G	Y	AT5G65640.1	LOC_Os03g04310.1
ZmbHLH120	GRMZM2G173372			1	7.996054	7.998214	IIIb	6	D	E non G	Y	AT5G65640.1	LOC_Os03g04310.1
ZmbHLH121	GRMZM2G128807		GRMZM5G857090	5	62.263888	62.265834	IIIb	4, 6	D	E non G	Y	AT5G65640.1	LOC_Os03g04310.2
ZmbHLH122	GRMZM2G173534	tac907.32		3	156.362362	156.365307	IIIb	4, 5, 6	D	E non G	Y	AT3G26744.1	LOC_Os01g70310.1
ZmbHLH123	GRMZM2G033356			4	23.423617	23.427358	IIIb	4, 5, 6	D	E non G	Y	AT3G26744.1	LOC_Os11g32100.1
ZmbHLH124	GRMZM2G172297			2	223.283172	223.285591	IIIb	4, 5, 6	D	E non G	Y	AT3G26744.1	LOC_Os11g32100.1
ZmbHLH125	GRMZM2G139372			4	237.951075	237.95461	IIIb	4, 7	D	E non G	Y	AT2G16910.1	LOC_Os02g02820.1
ZmbHLH126	GRMZM2G313756		GRMZM2G173521	10	108.358694	108.359794	IIIa+c		D	E non G	Y	AT2G28160.1	LOC_Os04g31290.1
ZmbHLH127	GRMZM2G173521		GRMZM2G313756	2	73.100275	73.101659	IIIa+c		D	E non G		AT2G28160.1	LOC_Os04g31290.1
ZmbHLH128	GRMZM2G107672			2	159.898523	159.899908	IIIa+c		D	E non G		AT2G28160.1	LOC_Os04g31290.1
ZmbHLH129	GRMZM2G119823			10	105.618184	105.620708	IIIa+c		D	E non G	Y	AT5G57150.4	LOC_Os04g23550.1
ZmbHLH130	GRMZM2G042895			10	76.586652	76.588863	IIIa+c		D	E non G	Y	AT5G57150.1	LOC_Os04g23550.1
ZmbHLH131	GRMZM2G042893			2	80.444662	80.447911	IIIa+c		D	E non G	Y	AT4G29930.2	LOC_Os04g23440.1
ZmbHLH132	GRMZM2G163233		GRMZM2G019999	2	204.467096	204.477531	IIIb		D	E non G	Y	AT3G42050.1	LOC_Os07g36470.1
ZmbHLH133	GRMZM2G150327			1	112.407261	112.409295	Orphan		D	BG		AT2G22760.1	LOC_Os10g01530.1
ZmbHLH134	GRMZM2G481280			6	6.914289	6.914735	Ib(1)		D	BG		AT1G49770.1	LOC_Os04g35010.1
ZmbHLH135	GRMZM2G147685			2	51.808541	51.813988	Ib(1)		D	BG	Y	AT1G49770.1	LOC_Os02g34320.1
ZmbHLH136	GRMZM2G475289		GRMZM2G171464	1	4.748612	4.750517	Ia		D	BG	Y	AT2G46810.1	LOC_Os03g03000.1
ZmbHLH137	GRMZM2G171464		GRMZM2G475289	9	153.843405	153.844854	Ia		D	BG	Y	AT3G24140.1	LOC_Os03g03000.1
ZmbHLH138	GRMZM2G417164			8	79.14171	79.142884	Ia	1, 2, 3	I	AG	Y	AT3G06120.1	LOC_Os05g51820.1
ZmbHLH139	GRMZM2G085751		GRMZM2G064638	4	224.62349	224.627055	Ia	1, 2	D	BG	Y	AT5G53210.1	LOC_Os02g15760.1
ZmbHLH140	GRMZM2G064638		GRMZM2G085751	5	140.933171	140.938202	Ia	1, 2	D	BG	Y	AT3G06120.1	LOC_Os02g15760.1
ZmbHLH141	GRMZM2G045109			9	67.071379	67.075188	Ia	1, 2	D	BG	Y	AT5G53210.1	LOC_Os06g33450.2
ZmbHLH142	GRMZM2G162450			8	75.990746	75.992414	Ia	1, 2, 3	D	BG	Y	AT3G24140.1	LOC_Os05g50900.1
ZmbHLH143	GRMZM2G111146		GRMZM2G091003	5	198.776715	198.780355	Ia	1, 2, 3	D	BG	Y	AT1G22490.1	LOC_Os02g46560.1
ZmbHLH144	GRMZM2G091003		GRMZM2G111146	4	157.480902	157.48364	Ia	1, 2, 3	D	BG	Y	AT3G24140.1	LOC_Os02g46560.1
ZmbHLH145	GRMZM2G086474			10	135.677365	135.680963	Ia	1	D	BG	Y	AT3G24140.1	LOC_Os02g46560.1
ZmbHLH146	GRMZM2G082586		GRMZM2G015666	7	128.707588	128.709815	Ia	1, 2, 3	D	BG	Y	AT1G72210.1	LOC_Os09g29360.1
ZmbHLH147	GRMZM2G015666		GRMZM2G082586	2	187.816227	187.818395	Ia	1, 2, 3	D	BG	Y	AT1G72210.1	LOC_Os09g29360.1

Table 3.3 (cont.)

bHLH No.	MaizeGDB #	Gene Name	Homeolog	Chr No.	Start (Mb)	End (Mb)	Subfamily	Motifs	Intron Pattern	DNA Binding	Expression	At Orthlog	Os Ortholog
ZmbHLH148	GRMZM2G006631		AC233899.1_FG002	1	19.099991	19.102005	Ia	1, 2, 3	D	BG	Y	AT1G72210.1	LOC_Os03g08930.1
ZmbHLH149	AC233899.1_FG002		GRMZM2G006631	9	149.664142	149.665651	Ia		D	BG		AT1G72210.1	LOC_Os03g08930.1
ZmbHLH150	GRMZM2G045883		GRMZM2G008691	5	208.89571	208.901046	Ia	1, 2, 3	D	BG	Y	AT1G72210.1	LOC_Os02g52190.1
ZmbHLH151	GRMZM2G008691		GRMZM2G045883	4	179.801255	179.803482	Ia	1, 2, 3	D	BG	Y	AT1G72210.1	LOC_Os02g52190.1
ZmbHLH152	GRMZM2G057413		AC193786.3_FG005	3	148.031503	148.032613	Ib2		D	BG	Y	AT3G56970.1	LOC_Os01g72370.1
ZmbHLH153	AC193786.3_FG005		GRMZM2G057413	8	157.922746	157.923803	Ib2		D	BG		AT3G56970.1	LOC_Os01g72370.1
ZmbHLH154	GRMZM2G137374			1	6.568931	6.570605	VIIIb		I	Non binder	Y	AT1G30670.1	LOC_Os12g31430.1
ZmbHLH155	GRMZM5G856837		GRMZM2G314882	5	69.447894	69.45052	Ivc	9	H	BG	Y	AT5G54680.1	LOC_Os02g02480.1
ZmbHLH156	GRMZM2G314882		GRMZM5G856837	4	240.035632	240.038589	Ivc	9	H	BG	Y	AT5G54680.1	LOC_Os02g02480.1
ZmbHLH157	GRMZM2G093744		GRMZM2G017586	10	83.592659	83.595951	Ivc	9	H	BG	Y	AT5G54680.1	LOC_Os08g04390.2
ZmbHLH158	GRMZM2G017586		GRMZM2G093744	4	35.575487	35.578048	Ivc	9	H	BG	Y	AT5G54680.1	LOC_Os08g04390.2
ZmbHLH159	GRMZM2G061906		GRMZM2G074725	6	151.125421	151.127557	Ivc	9	H	BG	Y	AT5G54680.1	LOC_Os05g38140.1
ZmbHLH160	GRMZM2G058451			7	155.341756	155.345811	Ivc	9	H	BG	Y	AT5G54680.1	LOC_Os07g35870.1
ZmbHLH161	GRMZM2G350312			1	65.657063	65.66029	Ivb	9	F	BG	Y	AT3G47640.1	LOC_Os03g26210.2
ZmbHLH162	GRMZM2G017145		GRMZM2G064938	7	167.4977	167.504924	Ivb		F	BG	Y		
ZmbHLH163	GRMZM2G133675			4	6.596882	6.601837	Ivb	9	F	BG	Y	AT3G47640.1	LOC_Os11g38870.2
ZmbHLH164	GRMZM2G114444			3	218.480031	218.485596	Ivb	9	F	BG	Y	AT3G19860.2	LOC_Os02g23823.1
ZmbHLH165	GRMZM2G021276			8	0.095823	0.098367	II		D	E non G	Y		LOC_Os01g18870.1
ZmbHLH166	AC233960.1_FG005			5	199.967465	199.969471	II		D	E non G		AT1G06170.1	LOC_Os04g51070.1
ZmbHLH167	GRMZM2G412430		GRMZM5G886266	1	67.849074	67.851238	XIV	28	I	Non binder	Y		LOC_Os03g27390.1
ZmbHLH168	GRMZM2G154029			9	120.745751	120.750044	XIV	28	I	Non binder	Y	AT1G27740.1	LOC_Os03g27390.1
ZmbHLH169	GRMZM2G454022		GRMZM2G163975	8	7.431648	7.435975	X	22	A	Non binder	Y	AT1G27660.1	LOC_Os01g01600.1
ZmbHLH170	GRMZM2G163975		GRMZM2G454022	3	31.453104	31.458374	X	22	A	Non binder	Y	AT1G27660.1	LOC_Os01g01600.1
ZmbHLH171	GRMZM2G440529			3	190.314296	190.317061	X	22	A	Non binder	Y	AT1G61660.1	LOC_Os01g57580.1
ZmbHLH172	GRMZM2G313058			6	156.762139	156.764429	X	22	A	Non binder	Y	AT1G61660.1	LOC_Os05g42180.1
ZmbHLH173	GRMZM2G155217		GRMZM5G838853	2	7.958755	7.962107	X	19, 22	B	Non binder	Y	AT1G61660.1	LOC_Os04g53990.1
ZmbHLH174	GRMZM2G004732		GRMZM2G460571	10	56.754859	56.757544	X	20, 22	A	Non binder	Y	AT1G61660.1	LOC_Os08g08160.1
ZmbHLH175	GRMZM2G074501		GRMZM2G009837	1	278.711464	278.714075	X	20, 22	B	Non E	Y	AT3G19500.1	LOC_Os03g55220.1
ZmbHLH176	GRMZM2G009837		GRMZM2G074501	5	6.508958	6.513515	X	20, 22	B	Non E	Y	AT3G19500.1	LOC_Os03g55220.1
ZmbHLH177	GRMZM2G155762			10	137.658752	137.664934	X	22	A	Non binder	Y	AT1G31050.1	LOC_Os04g47810.1
ZmbHLH178	GRMZM2G002627			5	194.108688	194.111934	X	20, 22	A	Non binder	Y	AT1G31050.1	LOC_Os02g45010.1
ZmbHLH179	GRMZM2G150729			1	42.011014	42.013283	X		A	Non binder	Y	AT1G27660.1	LOC_Os03g17130.1
ZmbHLH180	AC177924.2_FG001			2	70.717255	70.722062	VIIa+b (?)		B	Non binder		AT5G67110.1	LOC_Os02g54100.1
ZmbHLH181	GRMZM2G460472		GRMZM2G158281	2	140.657254	140.663126	XIII	25, 26, 27	D	Non E	Y	AT2G27230.1	LOC_Os11g06010.1
ZmbHLH182	GRMZM2G158281		GRMZM2G460472	4	212.314987	212.321049	XIII	25, 26, 27	D	Non E	Y	AT2G27230.1	LOC_Os11g06010.1
ZmbHLH183	GRMZM2G069408			3	143.052031	143.057877	XIII	25, 26, 27	D	Non E	Y	AT2G27230.1	LOC_Os12g06335.1
ZmbHLH184	GRMZM2G054564			10	25.567276	25.568319	IIId+e		D	Non E	Y		LOC_Os12g35070.1
ZmbHLH185	GRMZM5G863198			5	150.320122	150.321164	Ib2		D	E non G		AT4G20970.1	LOC_Os02g13670.1
ZmbHLH186	GRMZM2G169947			3	29.652219	29.652966	Ib2		D	E non G	Y	AT4G20970.1	LOC_Os01g01840.1
ZmbHLH187	GRMZM2G112629			3	29.612994	29.616601	Ib2			E non G	Y	AT1G10585.1	LOC_Os01g01870.1
ZmbHLH188	GRMZM2G301089			1	274.734582	274.736325	IIId+e		I	BG	Y	AT5G43650.1	LOC_Os03g53020.1
ZmbHLH189	GRMZM2G173862			8	172.396332	172.398396	IIId+e		D	E non G		AT5G43650.1	LOC_Os01g56690.1
ZmbHLH190	GRMZM2G109605			6	48.999704	49.001769	IIId+e		D	E non G		AT5G43650.1	LOC_Os03g53020.1
ZmbHLH191	GRMZM2G159456		AC149829.2_FG004	1	15.268455	15.269892	XV		D	Non binder	Y	AT3G28857.1	LOC_Os10g26410.1
ZmbHLH192	AC149829.2_FG004		GRMZM2G035156	9	151.129407	151.129907	XV		D	Non binder		AT3G28857.1	LOC_Os10g26410.1
ZmbHLH193	GRMZM2G042101		GRMZM2G036092	4	181.897625	181.898975	XV		D	Non binder	Y	AT3G28857.1	LOC_Os02g51320.1
ZmbHLH194	GRMZM2G036092		GRMZM2G042101	5	207.504983	207.506597	XV		D	Non binder	Y	AT1G26945.1	LOC_Os02g51320.1
ZmbHLH195	GRMZM2G072820	tac907.62	GRMZM2G414353	2	6.190524	6.191628	XV		D	Non binder		AT1G74500.1	LOC_Os04g54900.1
ZmbHLH196	GRMZM2G035156		AC149829.2_FG004	1	15.133983	15.136082	XV		D	Non binder	Y	AT3G47710.1	LOC_Os10g26460.1
ZmbHLH197	GRMZM2G389567		GRMZM2G100313	1	210.996014	210.996745	XIII		I	Non E	Y		LOC_Os08g34510.1

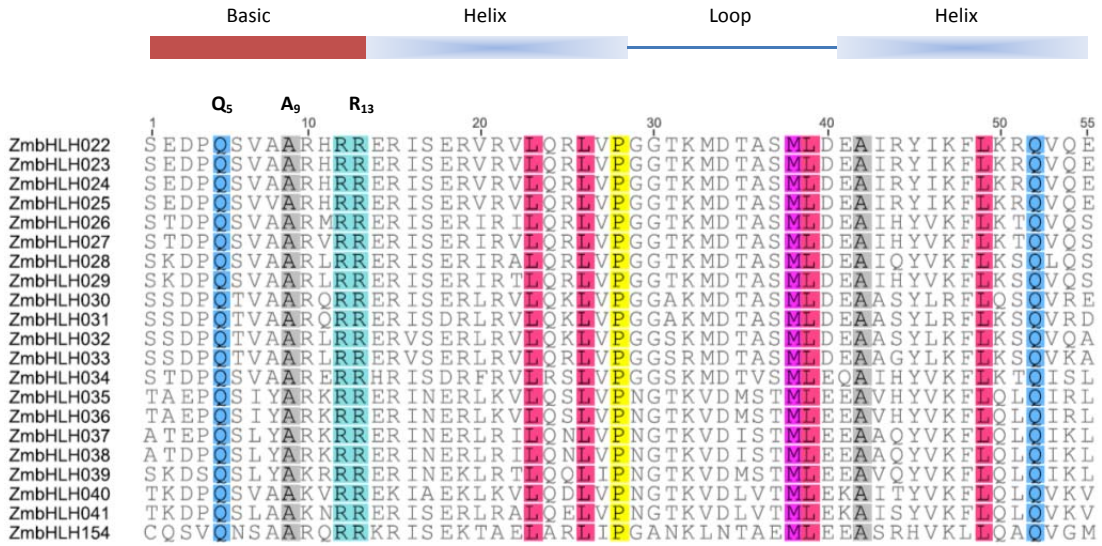


Fig 3.3: Alignment of maize bHLH domains containing the conserved Q_5-A_9-R_{13} motif. This motif has been proposed to recognize and bind to a novel DNA motif in plant genome (Pires and Dolan 2010).

	>5 basic amino acids			<5 basic amino acids		
	G-box	E non G	Non E binder	G-box	E non G	Non DNA binder
Maize	79	26	8	41	4	39
Rice	75	31	32	18	2	19
Arabidopsis	78	20	35	11	3	20

Table 3.4: Comparison of maize bHLHs with Arabidopsis and rice based upon its E-box, G-box, E-box (non G) and non DNA-binding features present in the N-terminal basic region. Arabidopsis and rice data obtained from (Carretero-Paulet et al. 2010).

1	ZmbHLH001	ERRR	41	ZmbHLH062	ERAR	81	ZmbHLH102	ERKR
2	ZmbHLH002	ERRR	42	ZmbHLH063	ERAR	82	ZmbHLH103	ERKR
3	ZmbHLH003	ERRR	43	ZmbHLH064	EKRR	83	ZmbHLH104	ERRR
4	ZmbHLH004	ERRR	44	ZmbHLH065	EKRR	84	ZmbHLH105	ERRR
5	ZmbHLH005	ERRR	45	ZmbHLH066	ERVR	85	ZmbHLH106	ERKR
6	ZmbHLH006	ERRR	46	ZmbHLH067	ERVR	86	ZmbHLH107	ERKR
7	ZmbHLH007	ERRR	47	ZmbHLH068	ERVR	87	ZmbHLH108	ERKR
8	ZmbHLH008	ER-RR	48	ZmbHLH069	ERVR	88	ZmbHLH109	ERQR
9	ZmbHLH009	EKRR	49	ZmbHLH070	ERVR	89	ZmbHLH110	ERQR
10	ZmbHLH010	EKRR	50	ZmbHLH071	ERER	90	ZmbHLH111	ERQR
11	ZmbHLH011	ERRR	51	ZmbHLH072	ERER	91	ZmbHLH112	ERQR
12	ZmbHLH012	ERRR	52	ZmbHLH073	ERER	92	ZmbHLH113	EHQR
13	ZmbHLH013	ERLR	53	ZmbHLH074	EQRR	93	ZmbHLH114	ERQR
14	ZmbHLH014	ERLR	54	ZmbHLH075	EQRR	94	ZmbHLH115	ERQR
15	ZmbHLH015	ERLR	55	ZmbHLH076	EQRR	95	ZmbHLH116	ERQR
16	ZmbHLH016	ERLR	56	ZmbHLH077	ERKR	96	ZmbHLH117	ERQR
17	ZmbHLH017	ERLR	57	ZmbHLH078	ERKR	97	ZmbHLH118	ERQR
18	ZmbHLH018	ERLR	58	ZmbHLH079	ERRR	98	ZmbHLH119	ERRR
19	ZmbHLH019	ERLR	59	ZmbHLH080	ERRR	99	ZmbHLH120	ERRR
20	ZmbHLH020	ERLR	60	ZmbHLH081	ERKR	100	ZmbHLH121	ERRR
21	ZmbHLH042	ERVR	61	ZmbHLH082	ERKR	101	ZmbHLH122	ERRR
22	ZmbHLH043	ERVR	62	ZmbHLH083	ERKR	102	ZmbHLH123	ERRR
23	ZmbHLH044	ERVR	63	ZmbHLH084	ERKR	103	ZmbHLH124	ERRR
24	ZmbHLH045	ERVR	64	ZmbHLH085	ERRR	104	ZmbHLH125	ERKR
25	ZmbHLH046	ERVR	65	ZmbHLH086	ERRR	105	ZmbHLH126	ERKR
26	ZmbHLH047	ERVR	66	ZmbHLH087	ERSR	106	ZmbHLH127	ERKR
27	ZmbHLH048	ERVR	67	ZmbHLH088	ERRR	107	ZmbHLH128	ERKR
28	ZmbHLH049	ERVR	68	ZmbHLH089	ERRR	108	ZmbHLH129	ERDR
29	ZmbHLH050	ERVR	69	ZmbHLH090	ERKR	109	ZmbHLH130	ERDR
30	ZmbHLH051	ERVR	70	ZmbHLH091	ERKR	110	ZmbHLH131	ERTR
31	ZmbHLH052	ERLR	71	ZmbHLH092	ERKR	111	ZmbHLH132	ERKR
32	ZmbHLH053	ERLR	72	ZmbHLH093	EKRR	112	ZmbHLH133	ERMR
33	ZmbHLH054	ERLR	73	ZmbHLH094	ERRR	113	ZmbHLH134	ERKR
34	ZmbHLH055	ERLR	74	ZmbHLH095	ERRR	114	ZmbHLH135	ERER
35	ZmbHLH056	ERVR	75	ZmbHLH096	ERRR	115	ZmbHLH136	ERNR
36	ZmbHLH057	ERVR	76	ZmbHLH097	EKRR	116	ZmbHLH137	ERNR
37	ZmbHLH058	ERFR	77	ZmbHLH098	ERKR	117	ZmbHLH138	ERNR
38	ZmbHLH059	ERFR	78	ZmbHLH099	ERKR	118	ZmbHLH139	ERNR
39	ZmbHLH060	ERAR	79	ZmbHLH100	ERKR	119	ZmbHLH140	ERNR
40	ZmbHLH061	ERAR	80	ZmbHLH101	ERKR	120	ZmbHLH141	ERNR

Table 3.5: The list of 150 ZmbHLH proteins containing conserved E-box binding motif (E_9-R/K_{12}). The sequences indicate the actual motif present.

Table 3.5 (cont.)

121	ZmbHLH142	ERNR
122	ZmbHLH143	ERNR
123	ZmbHLH144	ERNR
124	ZmbHLH145	ERNR
125	ZmbHLH146	ERNR
126	ZmbHLH147	ERNR
127	ZmbHLH148	ERNR
128	ZmbHLH149	ERNR
129	ZmbHLH150	ERNR
130	ZmbHLH151	ERNR
131	ZmbHLH152	ERDR
132	ZmbHLH153	ERGR
133	ZmbHLH155	EKMR
134	ZmbHLH156	EKMR
135	ZmbHLH157	EKMR
136	ZmbHLH158	ERMR
137	ZmbHLH159	EKLR
138	ZmbHLH160	EKIR
139	ZmbHLH161	EKLK
140	ZmbHLH162	EKRK
141	ZmbHLH163	EKLK
142	ZmbHLH164	EKMR
143	ZmbHLH165	EKQR
144	ZmbHLH166	ERER
145	ZmbHLH185	ERNR
146	ZmbHLH186	ERRR
147	ZmbHLH187	EKNR
148	ZmbHLH188	ERQR
149	ZmbHLH189	ERLR
150	ZmbHLH190	ERLR

1	ZmbHLH001	HNLSERRRR	41	ZmbHLH063	HSLAERARR	81	ZmbHLH111	HVESERQRR
2	ZmbHLH002	HNLSERRRR	42	ZmbHLH064	HNLTEKRRR	82	ZmbHLH112	HVEAERQRR
3	ZmbHLH003	HNLSERRRR	43	ZmbHLH065	HNLTEKRRR	83	ZmbHLH113	HVEAEHQRR
4	ZmbHLH004	HNLSERRRR	44	ZmbHLH074	HSATEQRRR	84	ZmbHLH114	HVEAERQRR
5	ZmbHLH005	HNLSERRRR	45	ZmbHLH075	HSATEQRRR	85	ZmbHLH115	HVEAERQRR
6	ZmbHLH006	HNLSERRRR	46	ZmbHLH076	HSATEQRRR	86	ZmbHLH116	HVEAERQRR
7	ZmbHLH007	HNLSERRRR	47	ZmbHLH077	HMISERKRR	87	ZmbHLH117	HVEAERQRR
8	ZmbHLH009	HNLSEKRRR	48	ZmbHLH078	HMLSERKRR	88	ZmbHLH118	HVEAERQRR
9	ZmbHLH010	HNLSEKRRR	49	ZmbHLH079	HMFSERRRR	89	ZmbHLH133	HIVAERMRR
10	ZmbHLH011	HNFSERRRR	50	ZmbHLH080	HMISERRRR	90	ZmbHLH134	HIVAERKRR
11	ZmbHLH012	HNFSERRRR	51	ZmbHLH081	HMFSERKRR	91	ZmbHLH135	HIFTERERR
12	ZmbHLH013	HSIAERLRR	52	ZmbHLH082	HMFSERKRR	92	ZmbHLH136	HIAVERNRR
13	ZmbHLH014	HSIAERLRR	53	ZmbHLH083	HVLSERKRR	93	ZmbHLH137	HIAVERNRR
14	ZmbHLH015	HSIAERLRR	54	ZmbHLH084	HVRSERKRR	94	ZmbHLH138	HIAVERNRR
15	ZmbHLH016	HSIAERLRR	55	ZmbHLH085	HSEAERRRR	95	ZmbHLH139	HITVERNRR
16	ZmbHLH017	HSIAERLRR	56	ZmbHLH086	HSEAERRRR	96	ZmbHLH140	HVAVERNRR
17	ZmbHLH018	HSIAERLRR	57	ZmbHLH087	HSEAERSRR	97	ZmbHLH141	HITVERNRR
18	ZmbHLH019	HSIAERLRR	58	ZmbHLH088	HSEAERRRR	98	ZmbHLH142	HIAVERNRR
19	ZmbHLH020	HSIAERLRR	59	ZmbHLH089	HSEAERRRR	99	ZmbHLH143	HIAVERNRR
20	ZmbHLH042	HSLAERVRR	60	ZmbHLH090	HSEAERKRR	100	ZmbHLH144	HIAVERNRR
21	ZmbHLH043	HSLAERVRR	61	ZmbHLH091	HSEAERKRR	101	ZmbHLH145	HIAVERNRR
22	ZmbHLH044	HSLAERVRR	62	ZmbHLH092	HSEAERKRR	102	ZmbHLH146	HIAVERNRR
23	ZmbHLH045	HSLAERVRR	63	ZmbHLH093	HSAAEKRRR	103	ZmbHLH147	HIAVERNRR
24	ZmbHLH046	HSLAERVRR	64	ZmbHLH094	HSEAERRRR	104	ZmbHLH148	HIAVERNRR
25	ZmbHLH047	HSLAERVRR	65	ZmbHLH095	HSEAERRRR	105	ZmbHLH149	HIAVERNRR
26	ZmbHLH048	HSLAERVRR	66	ZmbHLH096	HSEAERRRR	106	ZmbHLH150	HIAVERNRR
27	ZmbHLH049	HSLAERVRR	67	ZmbHLH097	HREAEKRRR	107	ZmbHLH151	HIAVERNRR
28	ZmbHLH050	HSLAERVRR	68	ZmbHLH098	HVMSERKRR	108	ZmbHLH152	HNAYERDRR
29	ZmbHLH051	HSLAERVRR	69	ZmbHLH099	HVMLERKRR	109	ZmbHLH153	HNAYERGRR
30	ZmbHLH052	HSLAERLRR	70	ZmbHLH100	HIIAERKRR	110	ZmbHLH155	KACREKMRR
31	ZmbHLH053	HSLAERLRR	71	ZmbHLH101	HIIAERKRR	111	ZmbHLH156	KACREKMRR
32	ZmbHLH054	HSLAERLRR	72	ZmbHLH102	HIIAERKRR	112	ZmbHLH157	KASREKMRR
33	ZmbHLH055	HSLAERLRR	73	ZmbHLH103	HIVAERKRR	113	ZmbHLH158	KACRERMRR
34	ZmbHLH056	HSLAERVRR	74	ZmbHLH104	HIIAERRRR	114	ZmbHLH159	KACREKLRR
35	ZmbHLH057	HSLAERVRR	75	ZmbHLH105	HIIAERRRR	115	ZmbHLH160	KASREKIRR
36	ZmbHLH058	HSLAERFRR	76	ZmbHLH106	HIVAERKRR	116	ZmbHLH161	KAEREKLKR
37	ZmbHLH059	HSLAERFRR	77	ZmbHLH107	HILAERKRR	117	ZmbHLH162	KSEREKRKR
38	ZmbHLH060	HSLAERARR	78	ZmbHLH108	HILAERKRR	118	ZmbHLH163	KAEREKLKR
39	ZmbHLH061	HSLAERARR	79	ZmbHLH109	HVEAERQRR	119	ZmbHLH164	KADREKMRR
40	ZmbHLH062	HSLAERARR	80	ZmbHLH110	HVEAERQRR	120	ZmbHLH188	HMMRERQRR

Table 3.6: List of 120 ZmbHLH proteins with conserved G-box binding motif ($(H/K)_5$-E_9-$(R/K)_{12}$-R_{13}). The sequences indicate the actual motif sequence present.

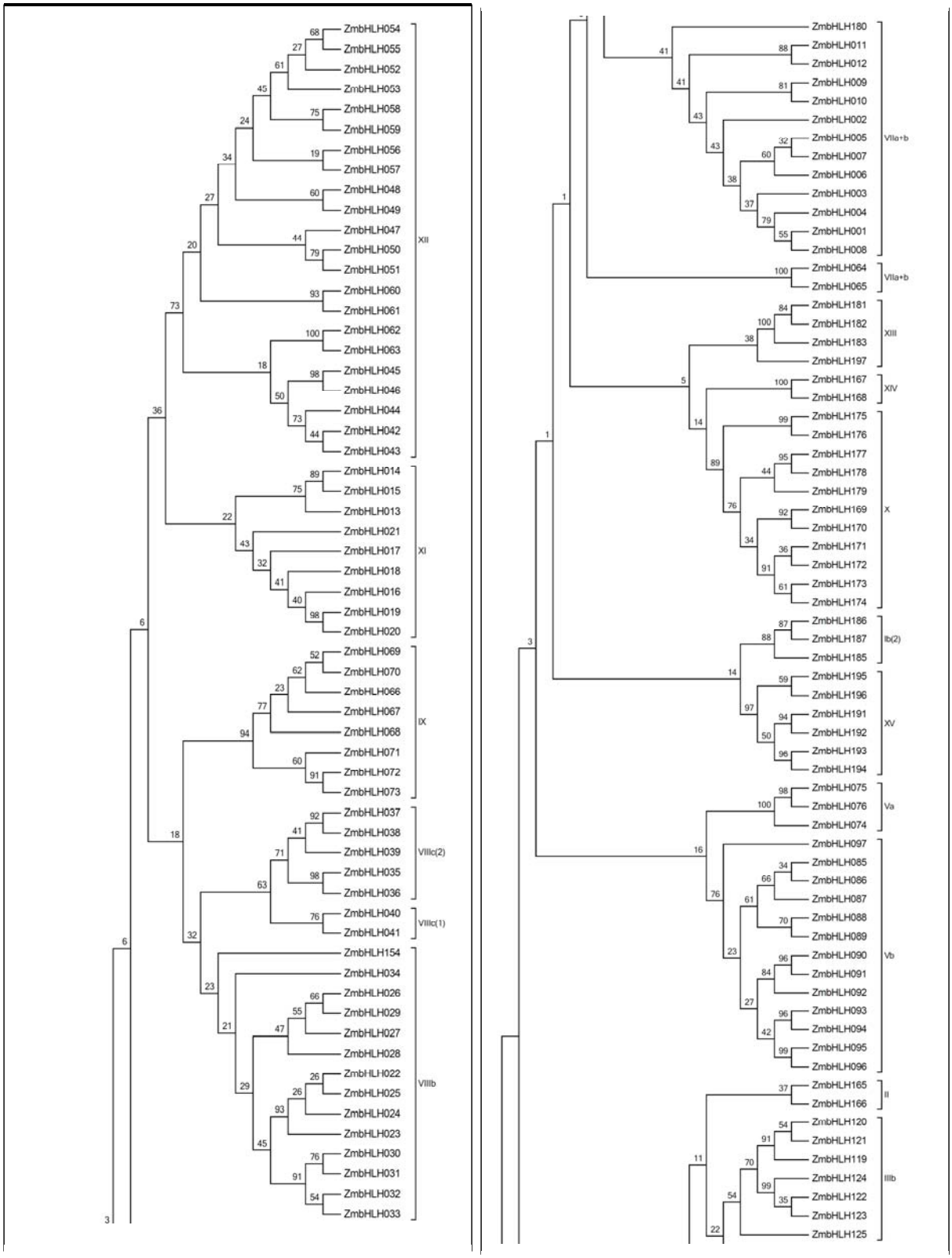
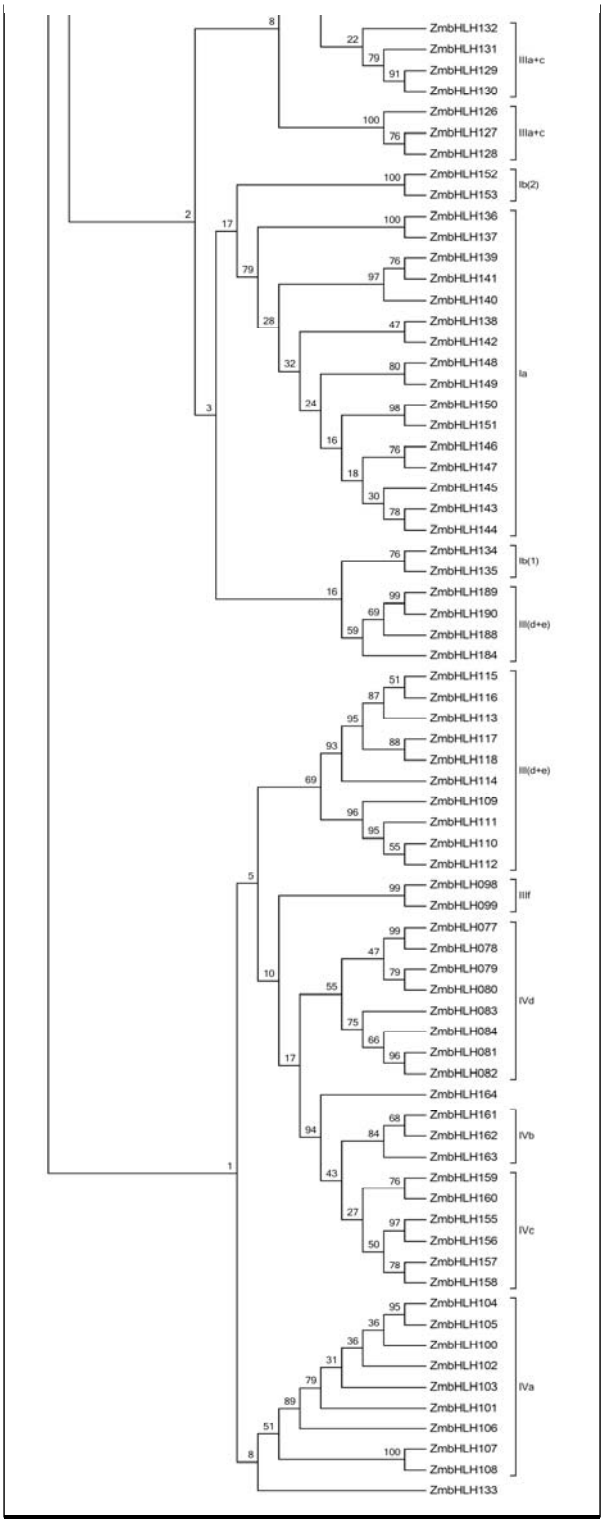


Fig 3.4: A Neighbor joining Phylogenetic tree of the 197 ZmbHLH proteins using the alignment of only bHLH domain region. The tree was created using MEGA v5.05 with 1000 bootstraps. ZmbHLH members were grouped into 26 subfamilies based upon the bootstrap support values (numbers at the nodes). Subfamily names are according to (Pires and Dolan 2010). ZmbHLH133 and ZmbHLH164 could not be grouped and are considered orphan.

Fig 3.4 (cont.)


ZmbHLH132
ZmbHLH131
ZmbHLH129
ZmbHLH130
IIIa+c
ZmbHLH126
ZmbHLH127
ZmbHLH128
IIIa+c
ZmbHLH152
ZmbHLH153
Ib(2)
ZmbHLH136
ZmbHLH137
ZmbHLH139
ZmbHLH141
ZmbHLH140
ZmbHLH138
ZmbHLH142
ZmbHLH148
ZmbHLH149
ZmbHLH150
ZmbHLH151
ZmbHLH146
ZmbHLH147
ZmbHLH145
ZmbHLH143
ZmbHLH144
Ia
ZmbHLH134
ZmbHLH135
Ib(1)
ZmbHLH189
ZmbHLH190
ZmbHLH188
ZmbHLH184
III(d+e)
ZmbHLH115
ZmbHLH116
ZmbHLH113
ZmbHLH117
ZmbHLH118
ZmbHLH114
ZmbHLH109
ZmbHLH111
ZmbHLH110
ZmbHLH112
III(d+e)
ZmbHLH098
ZmbHLH099
IIIf
ZmbHLH077
ZmbHLH078
ZmbHLH079
ZmbHLH080
ZmbHLH083
ZmbHLH084
ZmbHLH081
ZmbHLH082
IVd
ZmbHLH164
ZmbHLH161
ZmbHLH162
ZmbHLH163
IVb
ZmbHLH159
ZmbHLH160
ZmbHLH155
ZmbHLH156
ZmbHLH157
ZmbHLH158
IVc
ZmbHLH104
ZmbHLH105
ZmbHLH100
ZmbHLH102
ZmbHLH103
ZmbHLH101
ZmbHLH106
ZmbHLH107
ZmbHLH108
IVa
ZmbHLH133

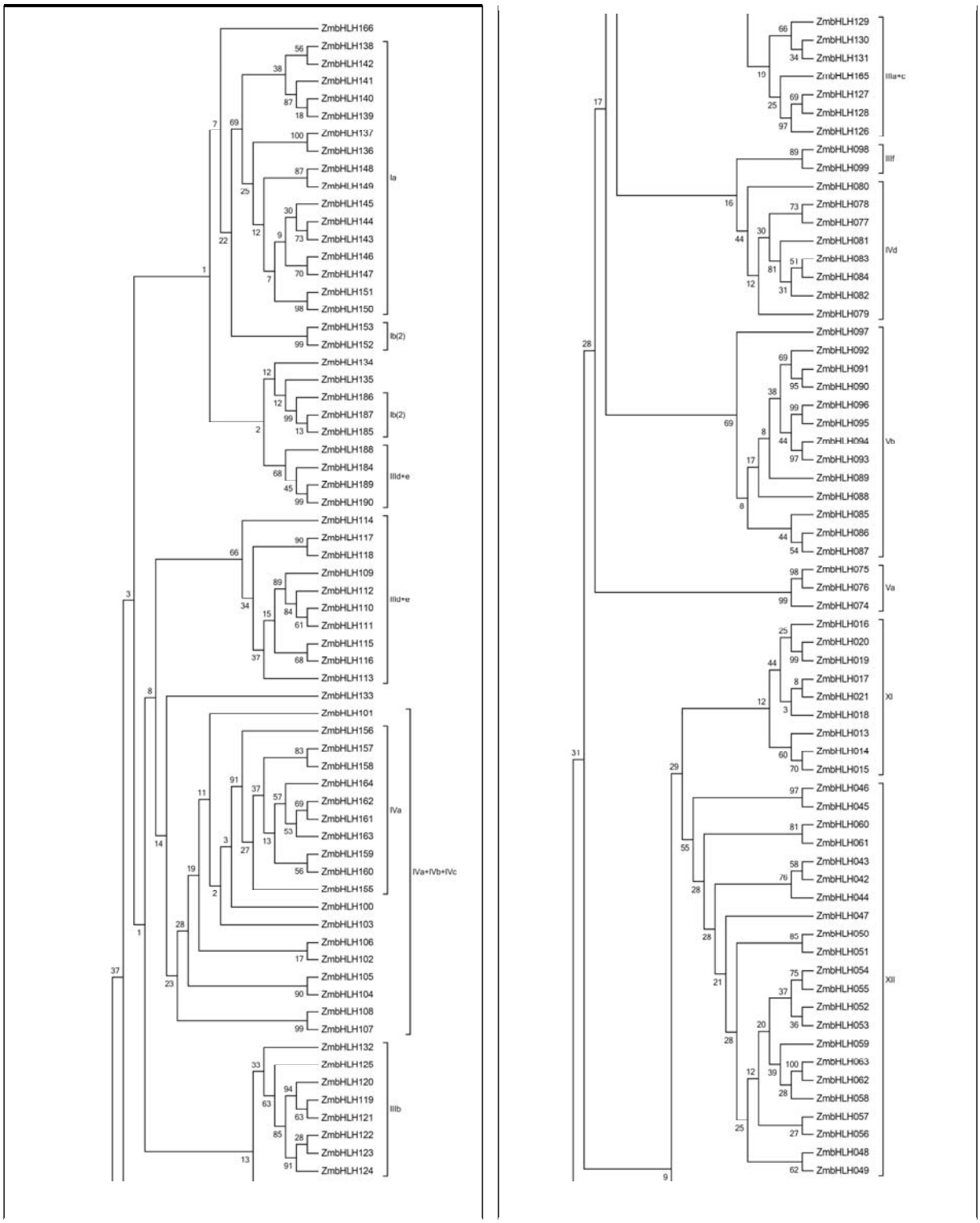
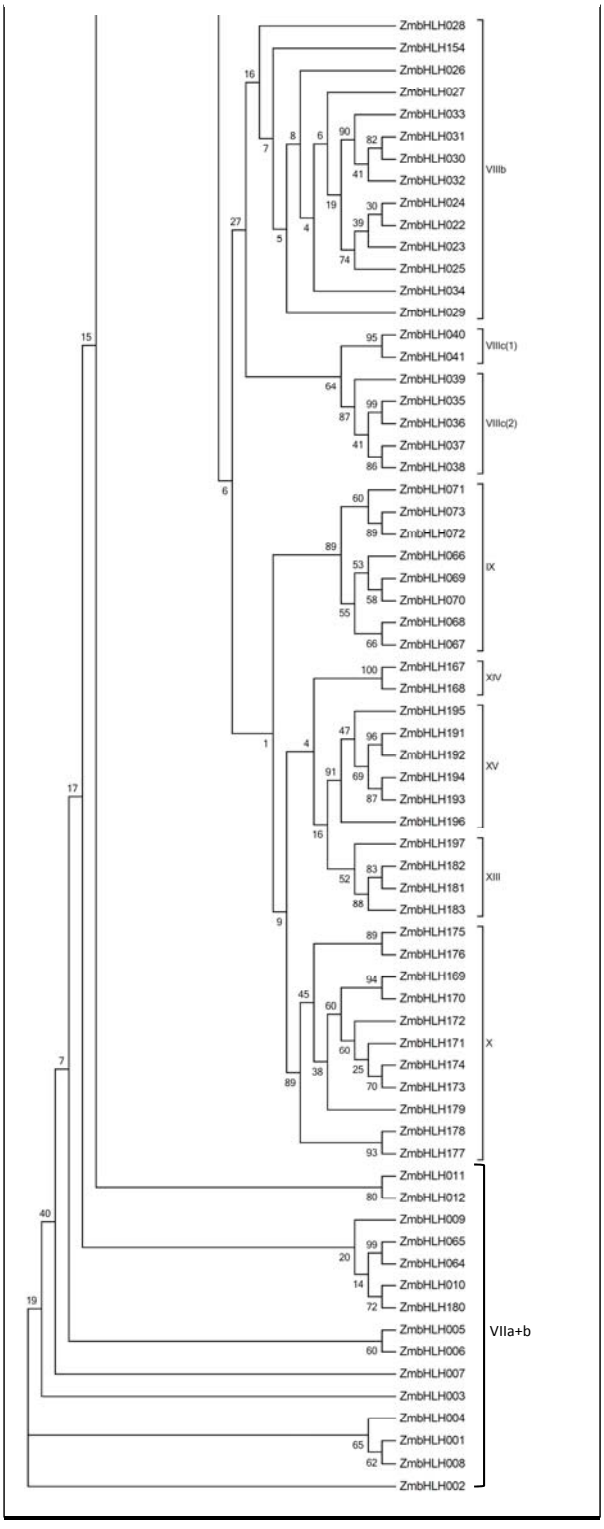


Fig 3.5: Maximum likelihood phylogenetic tree of ZmbHLH members created using 1000 bootstraps. Here, subfamily IVb and IVc could not be differentiated from Iva (marked as IVa+IVb+IVc). ZmbHLH134, ZmbHLH135 and ZmbHLH166 appeared as orphans.

Fig 3.5 (cont.)


ZmbHLH028
ZmbHLH154
ZmbHLH026
ZmbHLH027
ZmbHLH033
ZmbHLH031
ZmbHLH030
ZmbHLH032
ZmbHLH024
ZmbHLH022
ZmbHLH023
ZmbHLH025
ZmbHLH034
ZmbHLH029
VIIIb
ZmbHLH040
ZmbHLH041
VIIIc(1)
ZmbHLH039
ZmbHLH035
ZmbHLH036
ZmbHLH037
ZmbHLH038
VIIIc(2)
ZmbHLH071
ZmbHLH073
ZmbHLH072
ZmbHLH066
ZmbHLH069
ZmbHLH070
ZmbHLH068
ZmbHLH067
IX
ZmbHLH167
ZmbHLH168
XIV
ZmbHLH195
ZmbHLH191
ZmbHLH192
ZmbHLH194
ZmbHLH193
ZmbHLH196
XV
ZmbHLH197
ZmbHLH182
ZmbHLH181
ZmbHLH183
XIII
ZmbHLH175
ZmbHLH176
ZmbHLH169
ZmbHLH170
ZmbHLH172
ZmbHLH171
ZmbHLH174
ZmbHLH173
ZmbHLH179
ZmbHLH178
ZmbHLH177
X
ZmbHLH011
ZmbHLH012
ZmbHLH009
ZmbHLH065
ZmbHLH064
ZmbHLH010
ZmbHLH180
ZmbHLH005
ZmbHLH006
ZmbHLH007
ZmbHLH003
ZmbHLH004
ZmbHLH001
ZmbHLH008
ZmbHLH002
VIIa+b

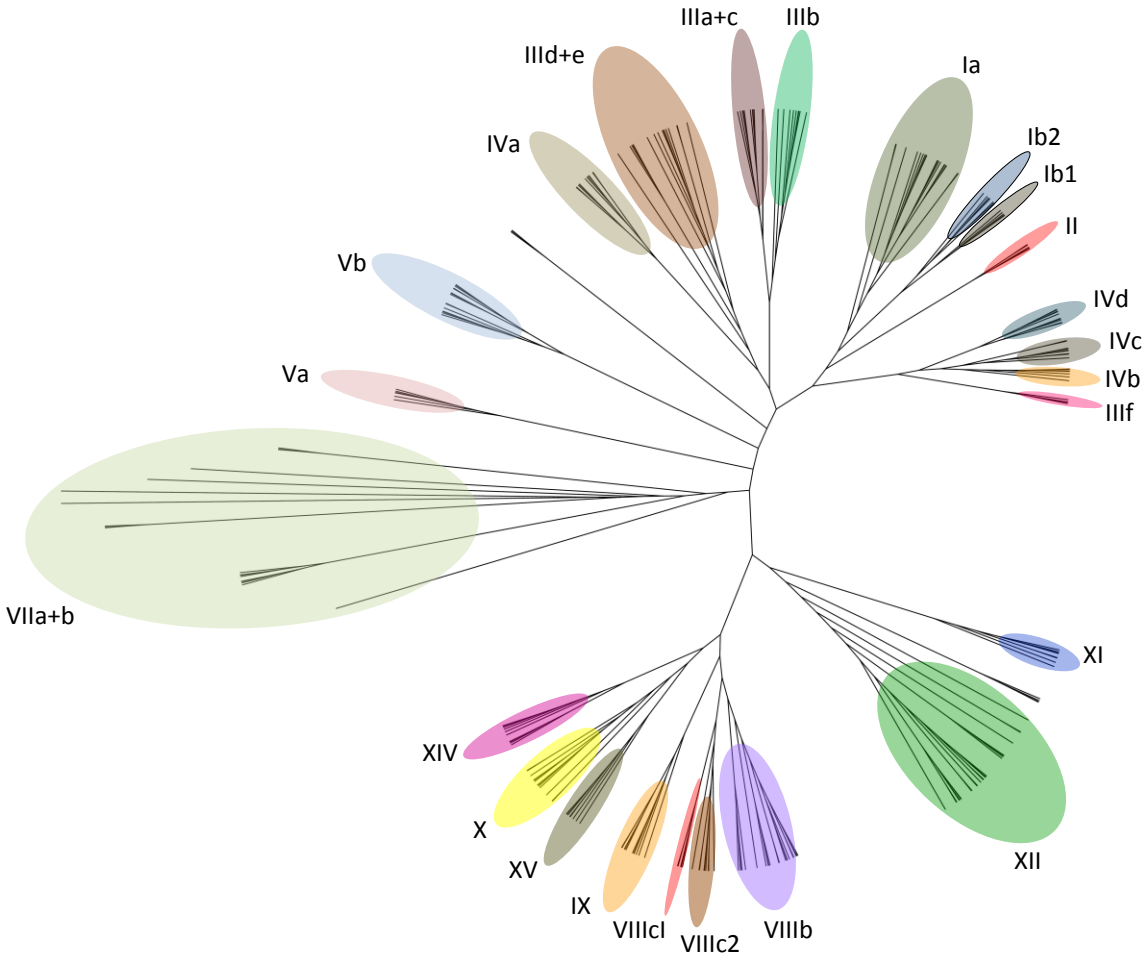


Fig 3.6: A cladogram showing the results of maximum likelihood phylogenetic analysis of ZmbHLH members and representative bHLH members of each subfamily in Arabidopsis, rice and *Physcomitrella*. Subfamily classification was also supported by non-bHLH motifs and intron pattern.

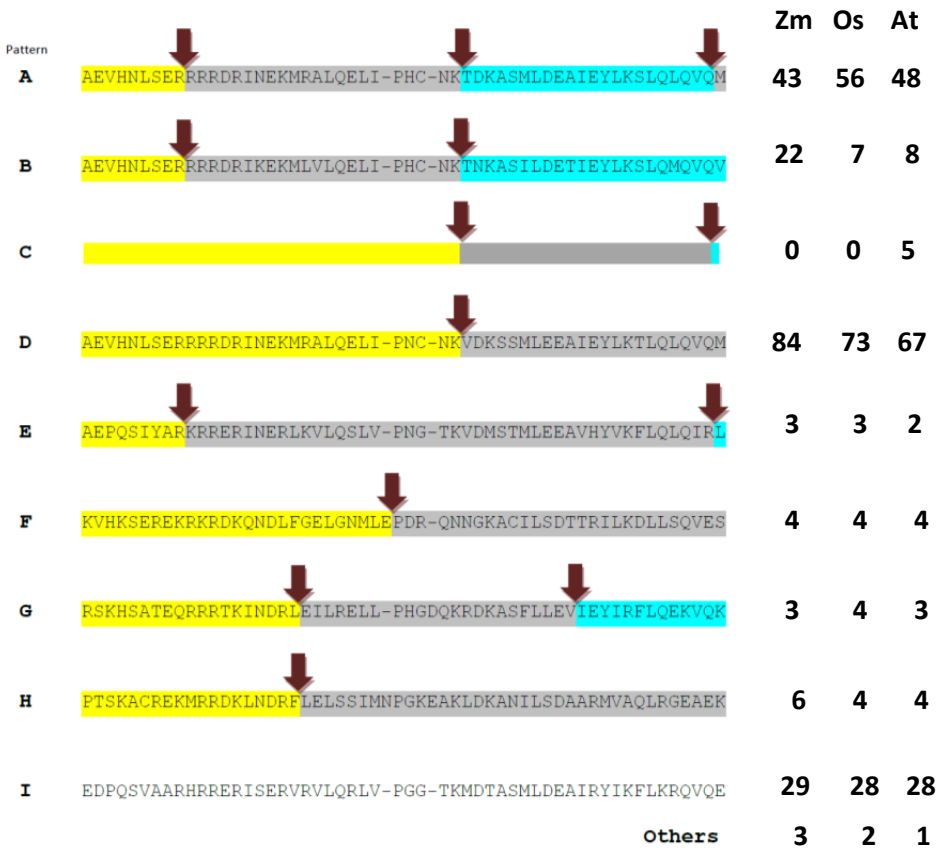


Fig 3.7: Intron pattern within bHLH domains. Alignment of 9 intron patterns with representative sequence. Locations of introns are indicated by red downward arrow. The numbers on the right represent the total number of genes with the given intron pattern. [Zm: *Zea Mays*; Os: *Oryza sativa*; At: *Arabidopsis thaliana*].

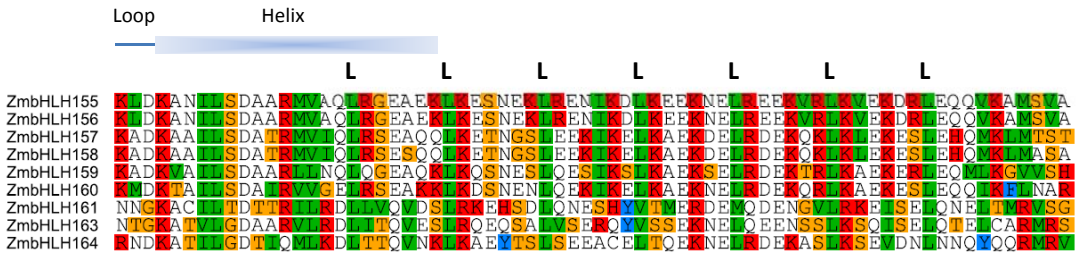


Fig 3.8: Alignment of ZmbHLHs having the leucine zipper motif (bHLH-Zip) present contiguous to the 2nd helical region of the bHLH domain. A conserved leucine (L) is present every 7 residues.

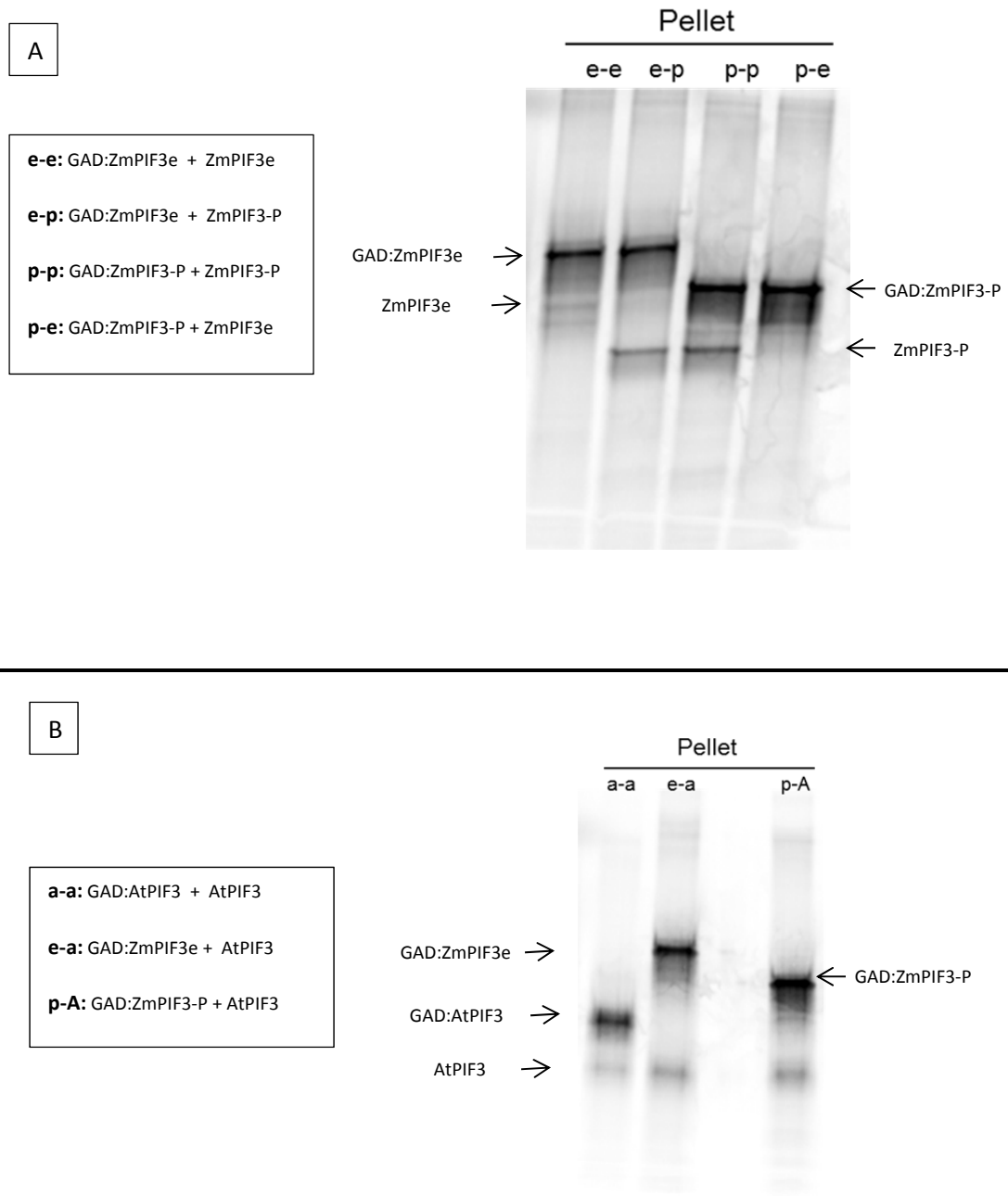


Fig 3.9: Co-immunoprecipitation showing homo- and hetero- dimer formation between ZmPIF3e and ZmPIF3-P. Phosphorimaging scans of SDS-PAGE- (A) Homo- and hetero-dimer between ZmPIF3e and ZmPIF3-P (B) hetero-dimer formation of AtPIF3 with maize PIF3e and PIF3-P.

Chapter 4

Differential gene Expression Analysis in Maize seedlings under shade conditions

Abstract

The shade avoidance syndrome has been a limiting factor in increasing yield for several crops. In this study, a replicated Affymetrix microarray experiment was used followed by statistical analysis to identify the genes differentially expressed in maize seedlings grown under simulated shade conditions. Functions of most of the predicted genes in maize are not known yet. It is hence not possible to assign a functional role of many differentially regulated maize genes that are potentially involved in the shade avoidance syndrome. Based on homologous genes in Arabidopsis or rice, I made an attempt to assign a putative role to many of the genes which could be potentially useful for future studies.

Introduction

Several crop species show the shade avoidance syndrome (SAS) typically as a response to close proximity planting. A high density of plants leads to a change in the quality of available light leading to low R (red) to FR (far-red) ratio, detected by phytochromes. SAS responses include elongation of stem, increased apical dominance, reduced branching, and decreased leaf area (Smith and Whitelam 1997) which help plants compete for light and space in a canopy. Under persistent shade conditions, plants accelerate flowering and show early seed production (Halliday et al. 1994). Physiological responses include a redistribution of auxin, increased ethylene production and an acceleration of flowering (Smith 1995; Morelli and Ruberti 2000; Vandenbussche et al. 2003). Light transmitted through any plant canopy contains attenuated red and blue wavelengths due to absorption by chlorophyll and carotenoids and utilization for photosynthesis. Thus, FR and green wavelengths are relatively enriched, leading to reduced R:FR ratio. The SAS has been shown to involve interaction between phytochrome and UV/blue light photoreceptor signaling pathways (Franklin 2008). Recently, it was shown that the abundant green light under canopy shade also informs plants of shade conditions and induces shade avoidance responses (Zhang et al. 2011). Under high R:FR conditions, shade avoidance responses are primarily suppressed by phyB mediated signaling (Nagatani et al. 1991; Somers et al. 1991) where phyB is mostly in the active Pfr form suppressing the excess elongation and early flowering (Whitelam and Devlin 1997). Various phytochrome mutant analyses in Arabidopsis revealed redundant roles of phyB, phyD and phyE (Aukerman et al. 1997; Devlin et

al. 1998; Devlin et al. 1999; Franklin et al. 2003). phyB and phyE have been shown to repress expression of a homeodomain leucine zipper (HD Zip) transcription factor *ATHB2*, a known promoter of SAS responses (Carabelli et al. 1996; Franklin et al. 2003). *PIL1* and *HFR1,* both bHLH transcription factor coding genes, are also early acting genes during SAS in Arabidopsis (Salter et al. 2003; Sessa et al. 2005).

Due to SAS responses, planting density has become a limiting factor in increasing crop yield. Interestingly, several maize cultivars with increased yields have been shown to possess higher tolerance for crowding (Duvick et al. 1997). The molecular response to high-density planting is not well characterized in maize, and there is potential for further yield improvement. There has been significant interest in using or modifying photoperception for crop improvement. Overexpression of phytochromes (phyA and phyB) has been repeatedly used to increase yield in densely-planted crops (Robson et al. 1996; Garg et al. 2006; Rao et al. 2011). Transgenic rice overexpressing the Arabidopsis *PHYA* gene was shown to have increases in yield, however it also showed unwanted pleiotropic effects such as dwarf plants with reduced internode length and diameter (Garg et al. 2006). Ectopic expression of Arabidopsis *PHYB* in potato also increased yield especially in densely planted crop and plants were shown to have higher stomatal conductance, transpiration rate and photosynthesis rates. (Thiele et al. 1999; Boccalandro et al. 2003; Schittenhelm et al. 2004). In a recent study, introduction and overexpression of the Arabidopsis *PHYB* gene in cotton showed ~35% increase in yield with pleiotropic effects like semi-dwarfism, decrease in apical dominance and increased boll size (Rao et al. 2011). Thus, manipulation of phytochrome mediated pathways also has immense potential in improving yield in maize. As there are fundamental differences in the monocot and

dicot phytochrome-mediated responses along with differences in the presence, sequence and expression of phytochrome gene family members, further understanding of the downstream components such as transcription factors and target genes during SAS in maize are required for future manipulation strategies. Here, I performed a microarray analysis for differential gene expression in maize seedlings grown under simulated shade conditions or control (dark) conditions. Previously, it has been shown that a pulse of far-red light immediately preceding the dark period (EOD-FR: end-of-day FR treatment) leads to increase in expression of *ATHB-2*, a known SAS response promoting gene, after 1 hour of the treatment (Carabelli et al. 1996; Franklin et al. 2003). It has also been shown that the EOD-FR treatment results in similar physiological responses to those observed under canopy shade or crowding (Fankhauser and Casal 2004). Therefore, EOD-FR treated maize seedlings were used for identification of differentially expressed genes. This study would provide information useful for biotechnological improvement of high yield crop varieties.

Materials and Method

EOD-FR treatment of maize B73 seedlings

Maize B73 seedlings were grown inside growth chambers under high intensity full spectrum light with 12 hour light / 12 hour dark daily cycles. At the end of 7 days after germination, immediately after the light period, the seedlings were treated with FR light for 1 hr followed by darkness. Corresponding control seedlings were kept continuously in darkness for the same hour after the light period. Seedlings were collected at the end of the 1hr treatment, and after 6hr of darkness following the treatment, and immediately frozen using liquid nitrogen. The frozen tissues were then stored at -80˚C. Tissue samples from each time point were pooled and four such sets were created to be used as biological replicates for each time point and corresponding control.

RNA extraction

RNA was extracted from 16 pooled tissue samples: four biological replicates each for dark 1 hr, FR 1 hr, dark 6 hr and FR 6 hr. RNA extractions were performed by Dr. Kankshita Swaminathan as per the method described (chapter 2).

Microarray Analysis

Microarray experiment was performed by Syngenta using the Syngenta maize Affymetrix chip. Affymetrix microarray data was normalized by RMAExpress (http://rmaexpress.bmbolstad.com/) software using standard settings (background adjust,

quantile normalization and median polish summarization). Normalized data was further analyzed by SAM v3.0 (Significance Analysis of Microarray) for Excel in two sets (1 hr time point and 6 hr time point) to identify differentially expressed genes. For SAM, the following settings were used: response type- two class unpaired; data type- array; test-statistics- Wilcoxon; regression method: standard; median centered; minimum fold change- 1.5; number of permutation- 100; and other settings were default.

Information about Affymetrix probeset ids with probe sequences and corresponding target genes was provided by Syngenta. However, this data was generated before the completion of the maize genome sequence, based on partial EST data. Due to the consequent recent updates on maize gene annotations, the Syngenta annotation was not used. Probe sequences for each significantly changing probeset id were used in a BLAST analysis to identify the target gene in the latest maize annotation release of working gene set (v5b.60). Probeset ids with only partial hits were not considered for further analysis. List of significantly changing genes are in Tables (4.1 and 4.2).

Multiple Sequence Alignment

Multiple sequence alignment of amino acids of DVL/RTFL members was created using MUSCLE in Geneious Pro v5.06 (Biomatters Ltd, Auckland, New Zealand)

Results and Discussion

Analysis of microarray data revealed upregulation and downregulation of several maize genes (Table 4.1 and 4.2). The criteria for the gene to be identified as differentially expressed were A) statistical significance for differential expression using SAM; B) minimum fold change of 1.5. Functions of most of the differentially regulated genes that were identified are currently not available from databases or the literature. For functional prediction, BLAST was performed using putative maize proteins against the Arabidopsis protein database (www.arabidopsis.org) and the nr database at NCBI. The top hit of the BLAST results for the corresponding maize genes are listed in Tables (4.1 and 4.2). Based upon information of Arabidopsis orthologs with known functions, some of the differentially expressed maize genes during simulated shade treatment could be assigned putative functions. The differentially expressed genes with putative functions are as follows:

Up-regulated genes

GRMZM2G057955_T01: The expression of this gene increased by 2.6 fold at 1 hr and 5 fold after 6 hours of EOD-FR treatment. The predicted gene product matched to AtETC1 (Arabidopsis enhancer of TRY (TRYPTYCHON) and CPC (CAPRICE), a small MYB protein. ETC1 has been shown to be involved in trichome and root hair formation (Kirik et al. 2004). ETC1 acts redundantly with TRY and CPC and inhibits cells to adopt trichome and non-root hair cells formation (Schellmann et al. 2002; Wada et al. 2002; Larkin et al. 2003). CPC has been shown to

inhibit expression of *GL2*, a gene involved in promoting hairless cells in root and thus CPC indirectly promotes root hair cell formation (Wada et al. 2002). Trichomes are hair-like appendages that develop from epidermal cells in aerial tissues (Werker 2000) and leaf trichomes have been shown to be primarily involved in protection against damage from herbivores (Bjorkman et al. 2008). Root hairs are known to aid in nutrient and water uptake from soil (Gilroy and Jones 2000). In terms of growth of a plant under shade condition where the plant display rapid elongation and early flowering, it is expected that the plant would require additional input of water and nutrients to cope up with the sudden surge of growth and additional root hairs would presumably help for this purpose. Inhibition of trichome formation is probably to allocate the resources for shade avoidance responses and in turn increasing susceptibility towards herbivory.

GRMZM2G167018_T01: The expression of this gene increased to 2 fold at 1 hour and 3 fold after 6 hours of FR treatment. The product of this gene shows resemblance to Arabidopsis NAC [(No Apical meristem (NAM)/Arabidopsis transcription activation factor (ATAF)/CUC (Cup-shaped cotyledon)] family of transcription factors. Members of this family have been shown to be involved in key developmental processes such as formation of shoot apical meristem and floral organs (Olsen et al. 2005). NAC family members NAM, CUC1 and CUC2 are involved in shoot apical meristem differentiation and development (Olsen et al. 2005). Several NAM/NAC family members have been shown to be involved in flowering as well (Hu et al. 2003). Increased apical dominance, shoot elongation and early flowering being major SAS responses, enhanced expression of NAC family of transcription factors as a result of EOD-FR light treatment seems plausible.

GRMZM2G103595_T01/GRMZM2G078164_T01: Expression of these maize genes also increased after FR light treatment to 1.5-1.8 fold. The predicted proteins coded by these genes are similar to DVL/RTFL (DEVIL/ROTUNDIFOLIA) family of proteins in Arabidopsis. DVL/RTFL family members are small proteins (50-60 amino acids) and have been suggested to act as signal polypeptides during plant development (Wen et al. 2004). In plants, more than 10 families of small peptides have been reported so far that act in plant development (Farrokhi et al. 2008). Arabidopsis and rice genome contains more than 20 and 25 DVL/RTFL members (Valdivia et al. 2012). DVL/RTFL family of proteins contains a conserved C-terminal domain with two highly conserved cysteine residues. The two maize DVL/RTFL members found here are longer (99 amino acids) compared to Arabidopsis members of this family, however they contain the conserved C-terminal domain (Fig 4.1). One of the Arabidopsis mutants, *dvl1-1D*, showed shorter but wider rosette leaves, short stature, tightly clustered inflorescence, shorter sepals and petals (Wen et al. 2004). Overexpression of DVL4 has been shown to change expression of at least 41 genes including several transcription factors (Larue et al. 2010). Interestingly, a *NAC1* gene was shown to genetically interact with a *DVL/ROT* family member (Larue et al. 2010). It has been suggested that DVL family members play role in various developmental process by mechanisms not known yet. It is possible that in response to FR light, these small polypeptides also act as signals for controlling various gene expressions involved in SAS.

GRMZM2G084958_T05: This gene was found to be upregulated after 6hr of the EOD-FR treatment. The Arabidopsis homolog for this gene encodes for PORA (NADPH:Protochlorophyllide oxidoreductase) that catalyzes the light-dependent reduction of protochlorophyllide *a* to chlorophyllide *a,* a metabolite in the biosynthesis of chlorophyll

(Masuda et al. 2003). Intriguingly, chlorophyll biosynthesis has been shown to be reduced under the shade avoidance syndrome (Smith and Whitelam 1997). Moreover, a reduction in *PORA* expression has also been seen in Arabidopsis seedlings when treated with FR light (Barnes et al. 1996). Although this microarray data is yet to be confirmed by another technique, it is possible that a different regulatory mechanism related to chlorophyll production exists in maize.

Downregulated Genes

GRMZM2G170692_T01: The expression of this gene was observed to go down by about 2 fold at the 6 hour time point after the EOD-FR exposure to maize seedlings. The maize genome also contains a homeolog (GRMZM2G063917) of this gene. The predicted protein shows homology to Arabidopsis PAL1 (Phenylanine Ammonia Lyase 1), an enzyme involved in anthocyanin and other phenylpropanoid biosynthesis. PAL catalyzes the first step in biosynthesis of phenylpropanoids (Wanner et al. 1995). Phenylpropanoids are primarily produced in response to biotic and abiotic stresses such as wounding, UV-irradiation, pollution and other non-favorable environmental conditions and provide protection through their antioxidant and free radical scavenging properties (Korkina 2007). Anthocyanin has been shown to reduce DNA damage by reducing the level of dimer formation in maize seedlings (Stapleton and Walbot 1994). In maize, it has been shown that anthocyanin production is photoinduced and is mediated by a coaction of phytochrome and UV-B light (Singh et al. 1999). Photoinduced accumulation of flavonoids is preceded by the induction of several enzymes involved in phenylpropanoid metabolism such as PAL, chalcone synthase, chalcone isomerase and

dihydroflavonol reductase (Hahlbrock and Scheel 1989; Kubasek et al. 1992). It has also been shown that the Pfr form of phytochrome (i.e. that formed under red light) induces anthocyanin production (Mohr and Drummherrel 1983). In rice, UV-B induced anthocyanin production was also found to be inhibited by the Pr form of phytochrome (i.e. that formed under FR light) (Reddy et al. 1994). Similarly, PAL activity was found to be reduced when tomato seedlings were treated with R followed by a FR pulse although longer exposure to FR increases PAL activity level (Lercari et al. 1982). Downregulation of the *Pal1* gene in maize seems consistent with EOD-FR treatment having decreased the expression due to an increase in Pr level compared to the dark control, where the level of phytochrome (Pfr) generated by the prior white light would still be high. Reduced production of phenylpropanoids has been shown to be both light and jasmonic acid dependent and it is likely part of resource reallocation for SAS responses (Cagnola et al. 2012).

GRMZM2G175076_T01: The Arabidopsis homolog of this gene encodes chalcone isomerase (CHI) which converts chalcone to flavonone, the step essential for biosynthesis for floral pigments and phenylpropanoids (Jez and Noel 2002). CHI and PAL1 function in the same biosynthetic pathway, thus the significance and regulation of maize *Chi* should be similar to *Pal1* mentioned above.

GRMZM2G114588_T01: This maize gene is similar to the isoflavone reductase coding gene in Arabidopsis. This gene is also involved in isoflavonoid biosynthesis, belonging to the class of phenylpropanoids as above.

GRMZM2G016241_T01: This maize gene is the homolog of *AtGSTU17* (Glutathione S Transferase 17), a member of a multifunctional dimeric enzyme superfamily. Plant GSTs have been shown to be involved in plant growth and development and bind to hormones such as auxin and cytokinin (Zettl et al. 1994; Gong et al. 2005). Several plant GSTs have been found to be induced by light and involved in phytochrome mediated signaling (Tepperman et al. 2001; Chen et al. 2007). *AtGSTU17* expression has been shown to be regulated by multiple photoreceptors, especially phyA (Tepperman et al. 2001) and possess roles in hypocotyl elongation, anthocyanin accumulation, and FR light mediated inhibition of greening (Jiang et al. 2010). In the study by (Jiang et al. 2010), AtGSTU17 expression was shown to be enhanced by FR light in Arabidopsis. In contrast, I found that the maize homolog is downregulated under EOD-FR treatment. GST is a big family of proteins in plants and it is possible that this maize GST possesses a different function and perhaps an antagonistic one. Further characterization of this gene is required.

GRMZM2G112538_T01: The Arabidopsis homolog of this gene encodes a member of PYR/PYL/RCAR (Pyrabactin resistance/PYR1-like/Regulatory components of ABA receptor). Members of this family (RCARs) bind to ABA and mediate inactivation of the ABI1 protein phosphatase 2C (Ma et al. 2009; Park et al. 2009; Nishimura et al. 2010). ABI1 is a protein phosphatase that plays a critical role as negative regulator of early abscisic acid responses (Hubbard et al. 2010). In this way, RCARs promote ABA responses. Interestingly, a low R:FR treatment has been shown to increase ABA level in *Brassica napus* (Kurepin et al. 2007) and under dark or FR treatment of Arabidopsis germinating seeds and seedlings (Weatherwax et al. 1996; Seo et al. 2006). ABA is known to function in the inhibition of growth. ABA responses

have been suggested to be mediated via multiple receptors including RCARs and GTGs; this is still an ongoing area of research (Muschietti and McCormick 2010). Based upon this information, it is likely that in the response to EOD-FR light, ABA responses mediated by RCARs are antagonistic to shade avoidance.

GRMZM2G106393_T01/GRMZM2G106445_T01/GRMZM2G106413_T01/ AC190933.3_FG004: These three genes belong to a wound-induced and stress-related family of proteins. Exact functions of these proteins are not yet known. However, it has been argued that plants create a balance in allocating the resources for competition (situations like shade avoidance) and defense against herbivory. Based on the circumstances the plant may adjust the allocations in response to growth conditions (Ballare 2011). It has been shown that expression of the shade avoidance syndrome is correlated with reduced expression of defense genes (Izaguirre et al. 2006; Moreno et al. 2009). Thus, the decrease of wound related genes in maize seedlings after EOD-FR treatment may just be the result of resource reallocation towards SAS responses as the seedling senses a potential shade condition.

Other genes: very little is known about the identity of the other maize genes identified as differentially expressed, and the corresponding Arabidopsis homologs including the bHLH transcription factor, rice *OSE3* and Myb DNA binding superfamily proteins. Further studies are required to characterize the function of these genes.

Conclusion

Maize seedlings showed a change in gene expression after EOD-FR treatment which mimics shade conditions. SAS responses include rapid elongation, enhanced apical dominance and resource allocation hence it is expected that the change in gene expression would be correlated towards these responses. Indeed, some of the upregulated genes indicate a role in elongation or potential signals for eliciting other SAS responses. Inhibition of several wound-response related genes indicate a resource allocation away from defense. Although the microarray data requires further verification, it provides a basic idea of critical genes involved in response to simulated shade condition and this information would be potentially useful for further studies related to SAS and yield increase in monocot crops.

Figures and Tables

A

Gene Id	Fold Change	Blast hit in At	Annotation
*GRMZM2G057955_T01	2.60	AT4G01060.1	MYB protein similar to enhancer of TRY and CPC
*GRMZM2G167018_T01	1.88	AT1G56010.2	NAM (No Apical Meristem) like
*GRMZM2G103595_T01	1.81	AT3G55515.1	Similar to AtDVL8/RTFL7
GRMZM2G167018_T01	1.65	AT1G56010.2	NAM (No Apical Meristem) like
*GRMZM2G078164_T01	1.56	AT2G39705.1	Similar to AtDVL11/RTFL8
GRMZM5G856777_T01	1.53	ATCG00280.1	Similar to PSBC (Photosystem II reaction center protein C)

B

Gene Id	Fold Change	BLAST Hit in At	Annotation
*GRMZM2G057955_T01	5.05	AT4G01060.1	MYB protein similar to enhancer of TRY and CPC
*GRMZM2G167018_T01	3.15	AT1G56010.2	Similar to NAM (No Apical Meristem)
GRMZM2G056093_T01	2.39	AT5G20190.1	Tetratricopeptide repeat (TPR)-like superfamily protein
GRMZM2G172214_T01	1.84	AT4G34120.1	CBS domain containing protein
*GRMZM2G078164_T01	1.63	AT2G39705.1	Similar to DVL11/RTFL8
GRMZM2G084958_T05	1.57	AT5G54190.1	Protochlorophyllide oxidoreductase A (PORA)
*GRMZM2G103595_T01	1.57	AT3G55515.1	Similar to DVL8/RTFL7
GRMZM2G033644_T01	1.55	AT3G06850.2	Dihydrolipoamide branched chain acyltransferase

Table 4.1: List of genes upregulated after (A) 1hr and (B) 6 hour of EOD-FR (End of day far-red light) treatment showing the fold change in expression and the top BLAST hit in Arabidopsis protein database. Annotation is based on the Arabidopsis ortholog. [At: *Arabidopsis thaliana;* * Genes appearing in both 1 hr and 6 hr list]

A

Gene Id	Fold Change	BLAST hit in At	Annotation
GRMZM2G046163_T01	2.01	AT4G02610.1; AT3G54640.1	Aldolase type TIM barrel family protein; Trp synthase alpha chain
GRMZM2G170692_T01	1.83	AT2G37040.1	AtPAL1 (Phe Ammonia Lyase 1)
GRMZM2G016241_T01	1.63	AT1G10370.1	Glutathione S transferase
GRMZM2G114588_T01	1.57	AT1G75280.1	Isoflavone reductase
GRMZM2G112538_T01	1.57	AT2G40330.1	PYL6 (Regulatory components of ABA receptor 9)
GRMZM2G175076_T01	1.56	AT5G05270.1	Chalcone isomerase

B

Gene Id	Fold Change	BLAST hit in At/Os	Annotation
GRMZM2G042895_T01	2.34	AT4G29930.2	bHLH superfamily protein
AC190933.3_FG004	2.27	AAM09541.1	Putative universal ethylene regulated stress protein USP1 (*Oryza sativa*)
AC149829.2_FG002	2.07	AF245482_1	OSE3 (*Oryza sativa*)
GRMZM2G106393_T01	1.89	AT4G10270.1	Wound induced protein (DUF3774)
GRMZM2G305362_T01	1.86	AT3G10040.1	Myb DNA binding SANT superfamily
GRMZM2G106445_T01	1.83	AT4G10270.1	Wound induced protein (DUF3774)
GRMZM2G106413_T01	1.73	AT4G10270.1	Wound induced protein (DUF3774)
GRMZM2G325693_T01	1.56	No hit	

Table 4.2: List of genes downregulated after (A) 1hr and (B) 6hr of EOD-FR (End of day far-red light treatment) showing the fold change in expression and the top BLAST hit in Arabidopsis protein database. Annotation is based on the Arabidopsis ortholog. [At: *Arabidopsis thaliana;* Os: *Oryza sativa*]

Fig 4.1: Alignment of Arabidopsis and maize DVL/RTFL family members showing the conserved C-terminal domain including the cysteine residues.

Chapter 5

General Discussion

Red and far-red light response pathways are crucial during plant development and for eliciting shade avoidance responses in plants. The maize genome contains only six phytochrome genes - A, B and C with two homeologs each (A1-A2, B1-B2, C1-C2). Here, in a yeast-two-hybrid screen using the C-terminal half of maize phyB1, a PIF3-like phyB1 interacting protein was found. PIF3 was one of the first PIF family members to be found in Arabidopsis and thus far has been associated with various light regulated functions including seedling development, anthocyanin biosynthesis and chloroplast development (Leivar and Quail 2011).

When *ZmPif3* cDNA was sequenced, two potential translational start sites, both in same ORF, were observed. Based on the similarity to the AtPIF3 N-terminal region, and an observation that the shorter ORF produced a protein that was nonfunctional in immunoprecipitation studies, the start site with the longer ORF was used for further studies and the full-length gene was referred as *ZmPif3e*. The maize genome also contains a homeolog of the *Pif3* gene identified in the two-hybrid screen. It is possible that this gene that may be functionally redundant, but also that it may be nonfunctional or have acquired a new or subfunction, since it was not isolated from the screen, but the *ZmPif3e* gene was recovered from many different unique clones. To investigate this possibility, the *ZmPif3e* homeolog was cloned from the maize cDNA library and referred as *ZmPif3-P*. Existence of duplicated phytochrome genes in maize may have led to differential function of each homeolog of the gene (Lynch and Force 2000). When checked with targeted

yeast two hybrid assays, both ZmPHYB1 and ZmPHYB2 showed interaction with ZmPIF3e that indicated potential redundant function of the homeologous proteins. In the case of ZmPIF3-P, it showed interaction with ZmPHYB1 however an interaction with ZmPHYB2 remained inconclusive. This suggests that both ZmPIF3e and ZmPIF3-P share biochemical functions transduced by ZmphyB1 – however subfunctionalization by means of differential expression of the two genes remains a distinct possibility.

Studies in Arabidopsis revealed that several phytochromes function redundantly, for example phyB, phyD and phyE in controlling flowering time and leaf development and phyA, phyB and phyE in seed germination (Li et al. 2011). Hence, it is expected that maize phytochromes may also have overlapping functions. Among other phytochromes, the C-terminal domain of ZmPHYA2 interacted with both PIF3e and PIF3-P (as did *ZmPhyB1* gene) however interactions of ZmPHYC1 again remained inconclusive. Three of the maize phytochromes (PHYB1, PHYB2 and PHYA2) cross-interacted with Arabidopsis PIF3, and in turn AtPHYB also showed potential interaction with maize PIF3 homeologs. This indicates a high level of functional conservation between the evolutionarily distant Arabidopsis and maize.

Full-length phytochrome apoproteins require a chromophore (phytochromobilin or phycocyanobilin) to form holophytochromes that can interconvert between Pfr and Pr forms under red or far-red light respectively. In an *in vitro* co-immunoprecipitation study, ZmphyB1 in its Pfr form showed binding to ZmPIF3e which is consistent with the finding in Arabidopsis. In preliminary experiments, other maize phytochromes were not observed to show any binding to ZmPIF3e. It would be an interesting finding if the two full length phyB homeologous proteins

show differential binding to PIF3e and PIF3-P; this would indicate different acquired subfunctions of the two homeologs of PHYB and PIF3.

PIFs belong to a bHLH transcription factor family. bHLH proteins are involved in a variety of developmental functions. Based upon information from Arabidopsis and rice, the maize genome was also expected to contain a large number of bHLH proteins. Using the latest maize genome sequence release (v5b.60), at least 197 bHLH members were found by bioinformatic analysis. Phylogenetic grouping of these bHLH proteins was also consistent with previous reports (Pires and Dolan 2010) and showed mostly conserved non-bHLH motifs and an intron pattern within the bHLH domain that was consistent with the phylogenetic clustering. This also suggests that bHLH families in maize potentially carry similar conserved overall functions as in Arabidopsis and rice. However, ~50% of the total bHLH genes were found to exist as homeolog pairs, each of which may have potentially evolved subfunctionalization or new acquired functions of the duplicated genes. This analysis also revealed at least 15 potential PIF family members, two of which I have cloned and demonstrated to interact with maize phytochromes. bHLH proteins form homo- or hetero –dimers and in a co-immunoprecipitation study, ZmPIF3-P showed homo-dimerization and hetero-dimerization with ZmPIF3e. A possible heterodimerization between the two proteins produced by the paired homeologous genes may be involved in a different regulatory function as it could change the DNA binding and/or protein binding features of the dimer. Moreover, both of the maize PIF3 homeologs also heterodimerized with Arabidopsis PIF3, indicating a conserved protein-protein interaction surface of the orthologous proteins between the two species.

The maize phytochrome-PIF system has not been studied in detail and there are several questions which remain to be answered. First, the ability of all the holophytochromes to physically interact with both PIF3 homeologs needs to be investigated. There are 15 members in the maize PIF family and their interaction potential with the 6 phytochromes needs to be determined. Most of the plant bHLHs recognize and bind to a G-box promoter motif on target genes. The promoters of maize genes can be screened for the presence of G-box motifs to identify potential PIF target genes. Phytochromes also form dimers and it still remains to be demonstrated that maize phytochromes can form homodimers, and the extent to which they can heterodimerize with homeologs and other phytochromes could be a potential mode of regulation for light mediated pathways. In Arabidopsis, phytochromes have been found to interact with other proteins such as NDPK2 and cryptochrome, a blue/UV light photoreceptor which indicated cross talk between different pathways. Orthologous genes from maize can be cloned and an interaction analysis could be performed to determine similar existence of cross-talk in maize.

In conclusion, this study suggests that the biochemical function and molecular structure of the phy-PIF3 system seems highly conserved between Arabidopsis and maize, despite sequence differences and evolutionary distance. However, the results also suggest that the duplicated phytochromes and *Pif3* genes resulting from the whole-genome duplication in the maize genome may have acquired different function or subfunctionalization. The maize genome also contains a large gene family of bHLH transcription factors which shows conserved sequence and thus potentially similar functions as already characterized in other plants such as Arabidopsis and rice. The existence of various homeologs and some shuffling of non-bHLH domains/motifs

within bHLH families again indicate the possibility of neo- and/or sub-functionalization of these genes in maize.

Bibliography

Al-Sady, B., E. A. Kikis, et al. (2008). "Mechanistic duality of transcription factor function in phytochrome signaling." Proc Natl Acad Sci U S A **105**(6): 2232-2237.
Al-Sady, B., W. Ni, et al. (2006). "Photoactivated phytochrome induces rapid PIF3 phosphorylation prior to proteasome-mediated degradation." Mol Cell **23**(3): 439-446.
Aravind, L. and E. V. Koonin (1999). "Gleaning non-trivial structural, functional and evolutionary information about proteins by iterative database searches." J Mol Biol **287**(5): 1023-1040.
Atchley, W. R. and W. M. Fitch (1997). "A natural classification of the basic helix-loop-helix class of transcription factors." Proceedings of the National Academy of Sciences of the United States of America **94**(10): 5172-5176.
Atchley, W. R., W. Terhalle, et al. (1999). "Positional dependence, cliques, and predictive motifs in the bHLH protein domain." Journal of molecular evolution **48**(5): 501-516.
Atchley, W. R. and J. Zhao (2007). "Molecular architecture of the DNA-binding region and its relationship to classification of basic helix-loop-helix proteins." Molecular biology and evolution **24**(1): 192-202.
Aukerman, M. J., M. Hirschfeld, et al. (1997). "A deletion in the PHYD gene of the Arabidopsis Wassilewskija ecotype defines a role for phytochrome D in red/far-red light sensing." Plant Cell **9**(8): 1317-1326.
Bae, G. and G. Choi (2008). "Decoding of light signals by plant phytochromes and their interacting proteins." Annual review of plant biology **59**: 281-311.
Bailey, P. C., C. Martin, et al. (2003). "Update on the basic helix-loop-helix transcription factor gene family in Arabidopsis thaliana." The Plant cell **15**(11): 2497-2502.
Ballare, C. L. (2011). "Jasmonate-induced defenses: a tale of intelligence, collaborators and rascals." Trends in plant science **16**(5): 249-257.
Barnes, S. A., N. K. Nishizawa, et al. (1996). "Far-red light blocks greening of Arabidopsis seedlings via a phytochrome A-mediated change in plastid development." The Plant cell **8**(4): 601-615.
Basu, D., K. Dehesh, et al. (2000). "Rice PHYC gene: structure, expression, map position and evolution." Plant Molecular Biology **44**(1): 27-42.
Bauer, D., A. Viczian, et al. (2004). "Constitutive photomorphogenesis 1 and multiple photoreceptors control degradation of phytochrome interacting factor 3, a transcription factor required for light signaling in Arabidopsis." Plant Cell **16**(6): 1433-1445.
Bjorkman, C., P. Dalin, et al. (2008). "Leaf trichome responses to herbivory in willows: induction, relaxation and costs." New Phytologist **179**(1): 176-184.
Blanc, G. and K. H. Wolfe (2004). "Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes." The Plant cell **16**(7): 1667-1678.
Boccalandro, H. E., E. L. Ploschuk, et al. (2003). "Increased phytochrome B alleviates density effects on tuber yield of field potato crops." Plant Physiol **133**(4): 1539-1546.
Boylan, M. T. and P. H. Quail (1996). "Are the phytochromes protein kinases?" Protoplasma **195**(1-4): 12-17.
Brownlie, P., T. Ceska, et al. (1997). "The crystal structure of an intact human Max-DNA complex: new insights into mechanisms of transcriptional control." Structure **5**(4): 509-520.
Cagnola, J. I., E. Ploschuk, et al. (2012). "Stem transcriptome reveals mechanisms to reduce the energetic cost of shade-avoidance responses in tomato." Plant physiology **160**(2): 1110-1119.
Carabelli, M., G. Morelli, et al. (1996). "Twilight-zone and canopy shade induction of the Athb-2 homeobox gene in green plants." Proc Natl Acad Sci U S A **93**(8): 3530-3535.

Carretero-Paulet, L., A. Galstyan, et al. (2010). "Genome-Wide Classification and Evolutionary Analysis of the bHLH Family of Transcription Factors in Arabidopsis, Poplar, Rice, Moss, and Algae." Plant Physiology **153**(3): 1398-1412.
Carretero-Paulet, L., A. Galstyan, et al. (2010). "Genome-wide classification and evolutionary analysis of the bHLH family of transcription factors in Arabidopsis, poplar, rice, moss, and algae." Plant Physiol **153**(3): 1398-1412.
Castelain, M., R. Le Hir, et al. (2012). "The non-DNA-binding bHLH transcription factor PRE3/bHLH135/ATBS1/TMO7 is involved in the regulation of light signaling pathway in Arabidopsis." Physiologia Plantarum **145**(3): 450-460.
Castillon, A., H. Shen, et al. (2007). "Phytochrome Interacting Factors: central players in phytochrome-mediated light signaling networks." Trends Plant Sci **12**(11): 514-521.
Chen, I. C., I. C. Huang, et al. (2007). "Glutathione S-transferase interacting with far-red insensitive 219 is involved in phytochrome A-mediated signaling in Arabidopsis." Plant physiology **143**(3): 1189-1202.
Chen, M., Y. Tao, et al. (2005). "Regulation of phytochrome B nuclear localization through light-dependent unmasking of nuclear-localization signals." Curr Biol **15**(7): 637-642.
Cheng, Y., G. Qin, et al. (2008). "NPY genes and AGC kinases define two key steps in auxin-mediated organogenesis in Arabidopsis." Proc Natl Acad Sci U S A **105**(52): 21017-21022.
Childs, K. L., F. R. Miller, et al. (1997). "The sorghum photoperiod sensitivity gene, Ma3, encodes a phytochrome B." Plant Physiol **113**(2): 611-619.
Chinnusamy, V., M. Ohta, et al. (2003). "ICE1: a regulator of cold-induced transcriptome and freezing tolerance in Arabidopsis." Genes Dev **17**(8): 1043-1054.
Chipman, D. M. and B. Shaanan (2001). "The ACT domain family." Curr Opin Struct Biol **11**(6): 694-700.
Choi, G., H. Yi, et al. (1999). "Phytochrome signalling is mediated through nucleoside diphosphate kinase 2." Nature **401**(6753): 610-613.
Christensen, A. H. and P. H. Quail (1989). "Structure and expression of a maize phytochrome-encoding gene." Gene **85**(2): 381-390.
Clack, T., S. Mathews, et al. (1994). "The phytochrome apoprotein family in Arabidopsis is encoded by five genes: the sequences and expression of PHYD and PHYE." Plant Mol Biol **25**(3): 413-427.
Crews, S. T. (1998). "Control of cell lineage-specific development and transcription by bHLH-PAS proteins." Genes & development **12**(5): 607-620.
Crozatier, M., D. Valle, et al. (1996). "Collier, a novel regulator of Drosophila head development, is expressed in a single mitotic domain." Current biology : CB **6**(6): 707-718.
Davidson, R. M., C. N. Hansey, et al. (2011). "Utility of RNA Sequencing for Analysis of Maize Reproductive Transcriptomes." Plant Gen. **4**(3): 191-203.
de Lucas, M., J. M. Daviere, et al. (2008). "A molecular framework for light and gibberellin control of cell elongation." Nature **451**(7177): 480-484.
Deng, X. W. and P. H. Quail (1999). "Signalling in light-controlled development." Semin Cell Dev Biol **10**(2): 121-129.
Devlin, P. F., S. R. Patel, et al. (1998). "Phytochrome E influences internode elongation and flowering time in Arabidopsis." Plant Cell **10**(9): 1479-1487.
Devlin, P. F., P. R. Robson, et al. (1999). "Phytochrome D acts in the shade-avoidance syndrome in Arabidopsis by controlling elongation growth and flowering time." Plant Physiol **119**(3): 909-915.
Doebley, J. F., B. S. Gaut, et al. (2006). "The molecular genetics of crop domestication." Cell **127**(7): 1309-1321.
Dubois, P. G., G. T. Olsefski, et al. (2010). "Physiological and genetic characterization of end-of-day far-red light response in maize seedlings." Plant Physiol **154**(1): 173-186.

Duek, P. D. and C. Fankhauser (2003). "HFR1, a putative bHLH transcription factor, mediates both phytochrome A and cryptochrome signalling." Plant J **34**(6): 827-836.
Duek, P. D. and C. Fankhauser (2005). "bHLH class transcription factors take centre stage in phytochrome signalling." Trends Plant Sci **10**(2): 51-54.
Eddy, S. R. (1998). "Profile hidden Markov models." Bioinformatics **14**(9): 755-763.
Essen, L. O., J. Mailliet, et al. (2008). "The structure of a complete phytochrome sensory module in the Pr ground state." Proc Natl Acad Sci U S A **105**(38): 14709-14714.
Fairchild, C. D., M. A. Schumaker, et al. (2000). "HFR1 encodes an atypical bHLH protein that acts in phytochrome A signal transduction." Genes Dev **14**(18): 2377-2391.
Fankhauser, C. and J. J. Casal (2004). "Phenotypic characterization of a photomorphogenic mutant." Plant J **39**(5): 747-760.
Fankhauser, C., K. C. Yeh, et al. (1999). "PKS1, a substrate phosphorylated by phytochrome that modulates light signaling in Arabidopsis." Science **284**(5419): 1539-1541.
Farrokhi, N., J. P. Whitelegge, et al. (2008). "Plant peptides and peptidomics." Plant biotechnology journal **6**(2): 105-134.
Fellner, M., L. A. Horton, et al. (2003). "Light interacts with auxin during leaf elongation and leaf angle development in young corn seedlings." Planta **216**(3): 366-376.
Feng, S., C. Martinez, et al. (2008). "Coordinated regulation of Arabidopsis thaliana development by light and gibberellins." Nature **451**(7177): 475-479.
Ferre-D'Amare, A. R., P. Pognonec, et al. (1994). "Structure and function of the b/HLH/Z domain of USF." The EMBO journal **13**(1): 180-189.
Ferre-D'Amare, A. R., G. C. Prendergast, et al. (1993). "Recognition by Max of its cognate DNA through a dimeric b/HLH/Z domain." Nature **363**(6424): 38-45.
Foster, K. R., F. R. Miller, et al. (1994). "Genetic Regulation of Development in Sorghum bicolor (VIII. Shoot Growth, Tillering, Flowering, Gibberellin Biosynthesis, and Phytochrome Levels Are Differentially Affected by Dosage of the ma3R Allele." Plant Physiol **105**(3): 941-948.
Franklin, K. A. (2008). "Shade avoidance." New Phytol **179**(4): 930-944.
Franklin, K. A., U. Praekelt, et al. (2003). "Phytochromes B, D, and E act redundantly to control multiple physiological responses in Arabidopsis." Plant Physiol **131**(3): 1340-1346.
Franklin, K. A. and P. H. Quail (2010). "Phytochrome functions in Arabidopsis development." J Exp Bot **61**(1): 11-24.
Franklin, K. A. and G. C. Whitelam (2004). "Light signals, phytochromes and cross-talk with other environmental cues." J Exp Bot **55**(395): 271-276.
Furutani, M., T. Kajiwara, et al. (2007). "The gene MACCHI-BOU 4/ENHANCER OF PINOID encodes a NPH3-like protein and reveals similarities between organogenesis and phototropism at the molecular level." Development **134**(21): 3849-3859.
Gallavotti, A., Q. Zhao, et al. (2004). "The role of barren stalk1 in the architecture of maize." Nature **432**(7017): 630-635.
Garg, A. K., R. J. Sawers, et al. (2006). "Light-regulated overexpression of an Arabidopsis phytochrome A gene in rice alters plant architecture and increases grain yield." Planta **223**(4): 627-636.
Gaut, B. S. (2001). "Patterns of chromosomal duplication in maize and their implications for comparative maps of the grasses." Genome Res **11**(1): 55-66.
Gaut, B. S. and J. F. Doebley (1997). "DNA sequence evidence for the segmental allotetraploid origin of maize." Proc Natl Acad Sci U S A **94**(13): 6809-6814.
Genoud, T., A. J. Buchala, et al. (2002). "Phytochrome signalling modulates the SA-perceptive pathway in Arabidopsis." Plant J **31**(1): 87-95.
Gilroy, S. and D. L. Jones (2000). "Through form to function: root hair development and nutrient uptake." Trends in Plant Science **5**(2): 56-60.

Gong, H., Y. Jiao, et al. (2005). "Expression of glutathione-S-transferase and its role in plant growth and development in vivo and shoot morphogenesis in vitro." Plant molecular biology **57**(1): 53-66.
Gremski, K., G. Ditta, et al. (2007). "The HECATE genes regulate female reproductive tract development in Arabidopsis thaliana." Development **134**(20): 3593-3601.
Hahlbrock, K. and D. Scheel (1989). "Physiology and Molecular-Biology of Phenylpropanoid Metabolism." Annual Review of Plant Physiology and Plant Molecular Biology **40**: 347-369.
Halliday, K. J., M. Koornneef, et al. (1994). "Phytochrome B and at Least One Other Phytochrome Mediate the Accelerated Flowering Response of Arabidopsis thaliana L. to Low Red/Far-Red Ratio." Plant Physiol **104**(4): 1311-1315.
Hatje, K., O. Keller, et al. (2011). "Cross-species protein sequence and gene structure prediction with fine-tuned Webscipio 2.0 and Scipio." BMC research notes **4**: 265.
Heim, M. A., M. Jakoby, et al. (2003). "The basic helix-loop-helix transcription factor family in plants: a genome-wide study of protein structure and functional diversity." Molecular biology and evolution **20**(5): 735-747.
Hiltbrunner, A., A. Tscheuschler, et al. (2006). "FHY1 and FHL act together to mediate nuclear accumulation of the phytochrome A photoreceptor." Plant Cell Physiol **47**(8): 1023-1034.
Hirschfeld, M., J. M. Tepperman, et al. (1998). "Coordination of phytochrome levels in phyB mutants of Arabidopsis as revealed by apoprotein-specific monoclonal antibodies." Genetics **149**(2): 523-535.
Holm, M., C. S. Hardtke, et al. (2001). "Identification of a structural motif that confers specific interaction with the WD40 repeat domain of Arabidopsis COP1." The EMBO journal **20**(1-2): 118-127.
Hornitschek, P., S. Lorrain, et al. (2009). "Inhibition of the shade avoidance response by formation of non-DNA binding bHLH heterodimers." EMBO J **28**(24): 3893-3902.
Hu, W., Y. Wang, et al. (2003). "Isolation, sequence analysis, and expression studies of florally expressed cDNAs in Arabidopsis." Plant molecular biology **53**(4): 545-563.
Hua, X., C. Yokoyama, et al. (1993). "SREBP-2, a second basic-helix-loop-helix-leucine zipper protein that stimulates transcription by binding to a sterol regulatory element." Proceedings of the National Academy of Sciences of the United States of America **90**(24): 11603-11607.
Hubbard, K. E., N. Nishimura, et al. (2010). "Early abscisic acid signal transduction mechanisms: newly discovered components and newly emerging questions." Genes & Development **24**(16): 1695-1708.
Huq, E., B. Al-Sady, et al. (2004). "Phytochrome-interacting factor 1 is a critical bHLH regulator of chlorophyll biosynthesis." Science **305**(5692): 1937-1941.
Huq, E., B. Al-Sady, et al. (2003). "Nuclear translocation of the photoreceptor phytochrome B is necessary for its biological function in seedling photomorphogenesis." Plant J **35**(5): 660-664.
Huq, E. and P. H. Quail (2002). "PIF4, a phytochrome-interacting bHLH factor, functions as a negative regulator of phytochrome B signaling in Arabidopsis." The EMBO journal **21**(10): 2441-2450.
Hyun, Y. and I. Lee (2006). "KIDARI, encoding a non-DNA Binding bHLH protein, represses light signal transduction in Arabidopsis thaliana." Plant molecular biology **61**(1-2): 283-296.
Izaguirre, M. M., C. A. Mazza, et al. (2006). "Remote sensing of future competitors: impacts on plant defenses." Proceedings of the National Academy of Sciences of the United States of America **103**(18): 7170-7174.
Jang, I. C., R. Henriques, et al. (2010). "Arabidopsis PHYTOCHROME INTERACTING FACTOR proteins promote phytochrome B polyubiquitination by COP1 E3 ligase in the nucleus." Plant Cell **22**(7): 2370-2383.
Jang, I. C., J. Y. Yang, et al. (2005). "HFR1 is targeted by COP1 E3 ligase for post-translational proteolysis during phytochrome A signaling." Genes & development **19**(5): 593-602.

Jez, J. M. and J. P. Noel (2002). "Reaction mechanism of chalcone isomerase. pH dependence, diffusion control, and product binding differences." The Journal of biological chemistry **277**(2): 1361-1369.
Jiang, H. W., M. J. Liu, et al. (2010). "A glutathione S-transferase regulated by light and hormones participates in the modulation of Arabidopsis seedling development." Plant physiology **154**(4): 1646-1658.
Jiao, Y., O. S. Lau, et al. (2007). "Light-regulated transcriptional networks in higher plants." Nat Rev Genet **8**(3): 217-230.
Kami, C., S. Lorrain, et al. (2010). "Light-regulated plant growth and development." Curr Top Dev Biol **91**: 29-66.
Kami, C., K. Mukougawa, et al. (2004). "Complementation of phytochrome chromophore-deficient Arabidopsis by expression of phycocyanobilin:ferredoxin oxidoreductase." Proceedings of the National Academy of Sciences of the United States of America **101**(4): 1099-1104.
Kanaoka, M. M., L. J. Pillitteri, et al. (2008). "SCREAM/ICE1 and SCREAM2 specify three cell-state transitional steps leading to arabidopsis stomatal differentiation." Plant Cell **20**(7): 1775-1785.
Kebrom, T. H., B. L. Burson, et al. (2006). "Phytochrome B represses Teosinte Branched1 expression and induces sorghum axillary bud outgrowth in response to light signals." Plant Physiol **140**(3): 1109-1117.
Khanna, R., E. Huq, et al. (2004). "A novel molecular recognition motif necessary for targeting photoactivated phytochrome signaling to specific basic helix-loop-helix transcription factors." The Plant cell **16**(11): 3033-3044.
Kim, W. Y., S. Fujiwara, et al. (2007). "ZEITLUPE is a circadian photoreceptor stabilized by GIGANTEA in blue light." Nature **449**(7160): 356-360.
Kircher, S., L. Kozma-Bognar, et al. (1999). "Light quality-dependent nuclear import of the plant photoreceptors phytochrome A and B." Plant Cell **11**(8): 1445-1456.
Kiribuchi, K., M. Sugimori, et al. (2004). "RERJ1, a jasmonic acid-responsive gene from rice, encodes a basic helix-loop-helix protein." Biochem Biophys Res Commun **325**(3): 857-863.
Kirik, V., M. Simon, et al. (2004). "The ENHANCER OF TRY AND CPC1 gene acts redundantly with TRIPTYCHON and CAPRICE in trichome and root hair cell patterning in Arabidopsis." Developmental biology **268**(2): 506-513.
Kong, S. G., D. S. Lee, et al. (2004). "Characterization of sunlight-grown transgenic rice plants expressing Arabidopsis phytochrome A." Molecular Breeding **14**(1): 35-45.
Korkina, L. G. (2007). "Phenylpropanoids as naturally occurring antioxidants: from plant defense to human health." Cellular and molecular biology **53**(1): 15-25.
Kubasek, W. L., B. W. Shirley, et al. (1992). "Regulation of Flavonoid Biosynthetic Genes in Germinating Arabidopsis Seedlings." The Plant cell **4**(10): 1229-1236.
Kunkel, T., V. Speth, et al. (1995). "In vivo characterization of phytochrome-phycocyanobilin adducts in yeast." The Journal of biological chemistry **270**(34): 20193-20200.
Kurepin, L. V., S. Shah, et al. (2007). "Light quality regulation of endogenous levels of auxin, abscisic acid and ethylene production in petioles and leaves of wild type and ACC deaminase transgenic Brassica napus seedlings." Plant Growth Regulation **52**(1): 53-60.
Lariguet, P., I. Schepens, et al. (2006). "PHYTOCHROME KINASE SUBSTRATE 1 is a phototropin 1 binding protein required for phototropism." Proceedings of the National Academy of Sciences of the United States of America **103**(26): 10134-10139.
Larionov, V., N. Kouprina, et al. (1996). "Specific cloning of human DNA as yeast artificial chromosomes by transformation-associated recombination." Proc Natl Acad Sci U S A **93**(1): 491-496.
Larkin, J. C., M. L. Brown, et al. (2003). "How do cells know what they want to be when they grow up? Lessons from epidermal patterning in Arabidopsis." Annual review of plant biology **54**: 403-430.

Larue, C. T., J. Q. Wen, et al. (2010). "Interactions between a NAC-Domain Transcription Factor and the Putative Small Protein Encoding DVL/ROT Gene Family." Plant Molecular Biology Reporter **28**(1): 162-168.
Ledent, V. and M. Vervoort (2001). "The basic helix-loop-helix protein family: comparative genomics and phylogenetic analysis." Genome research **11**(5): 754-770.
Ledent, V. and M. Vervoort (2001). "The basic helix-loop-helix protein family: comparative genomics and phylogenetic analysis." Genome Res **11**(5): 754-770.
Leivar, P., E. Monte, et al. (2008). "The Arabidopsis phytochrome-interacting factor PIF7, together with PIF3 and PIF4, regulates responses to prolonged red light by modulating phyB levels." The Plant cell **20**(2): 337-352.
Leivar, P., E. Monte, et al. (2008). "Multiple phytochrome-interacting bHLH transcription factors repress premature seedling photomorphogenesis in darkness." Curr Biol **18**(23): 1815-1823.
Leivar, P. and P. H. Quail (2011). "PIFs: pivotal components in a cellular signaling hub." Trends Plant Sci **16**(1): 19-28.
Leivar, P., J. M. Tepperman, et al. (2009). "Definition of early transcriptional circuitry involved in light-induced reversal of PIF-imposed repression of photomorphogenesis in young Arabidopsis seedlings." Plant Cell **21**(11): 3535-3553.
Lercari, B., F. Sodi, et al. (1982). "Phytochrome-Mediated Induction of Phenylalanine Ammonia-Lyase in the Cotyledons of Tomato (Lycopersicon-Esculentum Mill) Plants." Planta **156**(6): 546-552.
Li, J., G. Li, et al. (2011). "Phytochrome signaling mechanisms." Arabidopsis Book **9**: e0148.
Li, J., G. Li, et al. (2011). "Phytochrome signaling mechanisms." The Arabidopsis book / American Society of Plant Biologists **9**: e0148.
Li, L., K. Ljung, et al. (2012). "Linking photoreceptor excitation to changes in plant architecture." Genes Dev **26**(8): 785-790.
Li, X., X. Duan, et al. (2006). "Genome-wide analysis of basic/helix-loop-helix transcription factor family in rice and Arabidopsis." Plant Physiol **141**(4): 1167-1184.
Li, Y., X. Dai, et al. (2011). "NPY genes play an essential role in root gravitropic responses in Arabidopsis." Molecular plant **4**(1): 171-179.
Lorrain, S., T. Allen, et al. (2008). "Phytochrome-mediated inhibition of shade avoidance involves degradation of growth-promoting bHLH transcription factors." Plant J **53**(2): 312-323.
Lorrain, S., T. Allen, et al. (2008). "Phytochrome-mediated inhibition of shade avoidance involves degradation of growth-promoting bHLH transcription factors." The Plant journal : for cell and molecular biology **53**(2): 312-323.
Ludwig, S. R., L. F. Habera, et al. (1989). "Lc, a member of the maize R gene family responsible for tissue-specific anthocyanin production, encodes a protein similar to transcriptional activators and contains the myc-homology region." Proc Natl Acad Sci U S A **86**(18): 7092-7096.
Lynch, M. and A. Force (2000). "The probability of duplicate gene preservation by subfunctionalization." Genetics **154**(1): 459-473.
Ma, P. C., M. A. Rould, et al. (1994). "Crystal structure of MyoD bHLH domain-DNA complex: perspectives on DNA recognition and implications for transcriptional activation." Cell **77**(3): 451-459.
Ma, Y., I. Szostkiewicz, et al. (2009). "Regulators of PP2C phosphatase activity function as abscisic acid sensors." Science **324**(5930): 1064-1068.
Maddonni, G. A., M. E. Otegui, et al. (2002). "Maize leaves turn away from neighbors." Plant Physiol **130**(3): 1181-1189.
Martinez-Garcia, J. F., E. Huq, et al. (2000). "Direct targeting of light signals to a promoter element-bound transcription factor." Science **288**(5467): 859-863.

Massari, M. E. and C. Murre (2000). "Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms." Mol Cell Biol **20**(2): 429-440.
Masuda, T., N. Fusada, et al. (2003). "Functional analysis of isoforms of NADPH: protochlorophyllide oxidoreductase (POR), PORB and PORC, in Arabidopsis thaliana." Plant & cell physiology **44**(10): 963-974.
Mathews, S. and R. A. Sharrock (1996). "The phytochrome gene family in grasses (Poaceae): a phylogeny and evidence that grasses have a subset of the loci found in dicot angiosperms." Mol Biol Evol **13**(8): 1141-1150.
Mathews, S. and R. A. Sharrock (1997). "Phytochrome gene diversity." Plant Cell and Environment **20**(6): 666-671.
Matsushita, T., N. Mochizuki, et al. (2003). "Dimers of the N-terminal domain of phytochrome B are functional in the nucleus." Nature **424**(6948): 571-574.
Moglich, A., X. Yang, et al. (2010). "Structure and function of plant photoreceptors." Annu Rev Plant Biol **61**: 21-47.
Mohr, H. and H. Drummherrel (1983). "Coaction between Phytochrome and Blue Uv-Light in Anthocyanin Synthesis in Seedlings." Physiologia Plantarum **58**(3): 408-414.
Monte, E., B. Al-Sady, et al. (2007). "Out of the dark: how the PIFs are unmasking a dual temporal mechanism of phytochrome signalling." J Exp Bot **58**(12): 3125-3133.
Montgomery, B. L. and J. C. Lagarias (2002). "Phytochrome ancestry: sensors of bilins and light." Trends Plant Sci **7**(8): 357-366.
Morelli, G. and I. Ruberti (2000). "Shade avoidance responses. Driving auxin along lateral routes." Plant Physiol **122**(3): 621-626.
Moreno, J. E., Y. Tao, et al. (2009). "Ecological modulation of plant defense via phytochrome control of jasmonate sensitivity." Proceedings of the National Academy of Sciences of the United States of America **106**(12): 4935-4940.
Morgenstern, B. and W. R. Atchley (1999). "Evolution of bHLH transcription factors: modular evolution by domain shuffling?" Molecular biology and evolution **16**(12): 1654-1663.
Morohashi, K., M. Zhao, et al. (2007). "Participation of the Arabidopsis bHLH factor GL3 in trichome initiation regulatory events." Plant Physiol **145**(3): 736-746.
Murre, C., P. S. McCaw, et al. (1989). "A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins." Cell **56**(5): 777-783.
Muschietti, J. and S. McCormick (2010). "Abscisic acid (ABA) receptors: light at the end of the tunnel." F1000 biology reports **2**.
Nagatani, A. (2004). "Light-regulated nuclear localization of phytochromes." Curr Opin Plant Biol **7**(6): 708-711.
Nagatani, A. (2010). "Phytochrome: structural basis for its functions." Curr Opin Plant Biol **13**(5): 565-570.
Nagatani, A., J. Chory, et al. (1991). "Phytochrome-B Is Not Detectable in the Hy3 Mutant of Arabidopsis, Which Is Deficient in Responding to End-of-Day Far-Red Light Treatments." Plant and Cell Physiology **32**(7): 1119-1122.
Ni, M., J. M. Tepperman, et al. (1998). "PIF3, a phytochrome-interacting factor necessary for normal photoinduced signal transduction, is a novel basic helix-loop-helix protein." Cell **95**(9845368): 657-667.
Ni, M., J. M. Tepperman, et al. (1998). "PIF3, a phytochrome-interacting factor necessary for normal photoinduced signal transduction, is a novel basic helix-loop-helix protein." Cell **95**(5): 657-667.
Ni, M., J. M. Tepperman, et al. (1999). "Binding of phytochrome B to its nuclear signalling partner PIF3 is reversibly induced by light." Nature **400**(6746): 781-784.

Nishimura, N., A. Sarkeshik, et al. (2010). "PYR/PYL/RCAR family members are major in-vivo ABI1 protein phosphatase 2C-interacting proteins in Arabidopsis." The Plant journal : for cell and molecular biology **61**(2): 290-299.
Odronitz, F., H. Pillmann, et al. (2008). "WebScipio: an online tool for the determination of gene structures using protein sequences." BMC genomics **9**: 422.
Oh, E., J. Kim, et al. (2004). "PIL5, a phytochrome-interacting basic helix-loop-helix protein, is a key negative regulator of seed germination in Arabidopsis thaliana." The Plant cell **16**(11): 3045-3058.
Oh, E., S. Yamaguchi, et al. (2007). "PIL5, a phytochrome-interacting bHLH protein, regulates gibberellin responsiveness by binding directly to the GAI and RGA promoters in Arabidopsis seeds." Plant Cell **19**(4): 1192-1208.
Ohashi-Ito, K. and D. C. Bergmann (2007). "Regulation of the Arabidopsis root vascular initial population by LONESOME HIGHWAY." Development **134**(16): 2959-2968.
Oka, Y., T. Matsushita, et al. (2004). "Functional analysis of a 450-amino acid N-terminal fragment of phytochrome B in Arabidopsis." Plant Cell **16**(8): 2104-2116.
Olsen, A. N., H. A. Ernst, et al. (2005). "NAC transcription factors: structurally distinct, functionally diverse." Trends in Plant Science **10**(2): 79-87.
Osterlund, M. T., C. S. Hardtke, et al. (2000). "Targeted destabilization of HY5 during light-regulated development of Arabidopsis." Nature **405**(6785): 462-466.
Osugi, A., H. Itoh, et al. (2011). "Molecular dissection of the roles of phytochrome in photoperiodic flowering in rice." Plant Physiol **157**(3): 1128-1137.
Paik, I., S. Yang, et al. (2012). "Phytochrome regulates translation of mRNA in the cytosol." Proc Natl Acad Sci U S A **109**(4): 1335-1340.
Palagyi, A., K. Terecskei, et al. (2010). "Functional analysis of amino-terminal domains of the photoreceptor phytochrome B." Plant Physiol **153**(4): 1834-1845.
Paris, S., R. Longhi, et al. (2003). "Leucine-zipper-mediated homo- and hetero-dimerization of GIT family p95-ARF GTPase-activating protein, PIX-, paxillin-interacting proteins 1 and 2." The Biochemical journal **372**(Pt 2): 391-398.
Park, E., J. Kim, et al. (2004). "Degradation of phytochrome interacting factor 3 in phytochrome-mediated light signaling." Plant Cell Physiol **45**(8): 968-975.
Park, M. Y., M. S. Chung, et al. (2009). "Isolation and functional characterization of the Arabidopsis salt-tolerance 32 (AtSAT32) gene associated with salt tolerance and ABA signaling." Physiologia Plantarum **135**(4): 426-435.
Parks, B. M. and P. H. Quail (1991). "Phytochrome-Deficient hy1 and hy2 Long Hypocotyl Mutants of Arabidopsis Are Defective in Phytochrome Chromophore Biosynthesis." The Plant cell **3**(11): 1177-1186.
Payne, C. T., F. Zhang, et al. (2000). "GL3 encodes a bHLH protein that regulates trichome development in arabidopsis through interaction with GL1 and TTG1." Genetics **156**(3): 1349-1362.
Pires, N. and L. Dolan (2010). "Origin and diversification of basic-helix-loop-helix proteins in plants." Molecular biology and evolution **27**(4): 862-874.
Qian, Y., J. Zhang, et al. (2008). "DEC1, a basic helix-loop-helix transcription factor and a novel target gene of the p53 family, mediates p53-dependent premature senescence." The Journal of biological chemistry **283**(5): 2896-2905.
Quail, P. H. (1997). "An emerging molecular map of the phytochromes." Plant Cell and Environment **20**(6): 657-665.
Quail, P. H. (1997). "The phytochromes: a biochemical mechanism of signaling in sight?" Bioessays **19**(7): 571-579.
Quail, P. H. (2002). "Phytochrome photosensory signalling networks." Nat Rev Mol Cell Biol **3**(2): 85-93.

Quail, P. H., M. T. Boylan, et al. (1995). "Phytochromes: photosensory perception and signal transduction." Science **268**(5211): 675-680.
Rao, A. Q., M. Irfan, et al. (2011). "Overexpression of the phytochrome B gene from Arabidopsis thaliana increases plant growth and yield of cotton (Gossypium hirsutum)." J Zhejiang Univ Sci B **12**(4): 326-334.
Reddy, V. S., K. V. Goud, et al. (1994). "Ultraviolet-B-Responsive Anthocyanin Production in a Rice Cultivar Is Associated with a Specific Phase of Phenylalanine Ammonia-Lyase Biosynthesis." Plant Physiology **105**(4): 1059-1066.
Ren, J., L. Wen, et al. (2009). "DOG 1.0: illustrator of protein domain structures." Cell research **19**(2): 271-273.
Riechmann, J. L., J. Heard, et al. (2000). "Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes." Science **290**(5499): 2105-2110.
Rizzini, L., J. J. Favory, et al. (2011). "Perception of UV-B by the Arabidopsis UVR8 protein." Science **332**(6025): 103-106.
Robinson, K. A., J. I. Koepke, et al. (2000). "A network of yeast basic helix-loop-helix interactions." Nucleic acids research **28**(22): 4460-4466.
Robson, P. R., A. C. McCormac, et al. (1996). "Genetic engineering of harvest index in tobacco through overexpression of a phytochrome gene." Nat Biotechnol **14**(8): 995-998.
Rockwell, N. C. and J. C. Lagarias (2010). "A brief history of phytochromes." Chemphyschem **11**(6): 1172-1180.
Rockwell, N. C., Y.-S. Su, et al. (2006). "Phytochrome structure and signaling mechanisms." Annu Rev Plant Biol **57**(16669784): 837-858.
Roig-Villanova, I., J. Bou-Torrent, et al. (2007). "Interaction of shade avoidance and auxin responses: a role for two novel atypical bHLH proteins." The EMBO journal **26**(22): 4756-4767.
Rosler, J., I. Klein, et al. (2007). "Arabidopsis fhl/fhy1 double mutant reveals a distinct cytoplasmic action of phytochrome A." Proc Natl Acad Sci U S A **104**(25): 10737-10742.
Salter, M. G., K. A. Franklin, et al. (2003). "Gating of the rapid shade-avoidance response by the circadian clock in plants." Nature **426**(6967): 680-683.
Sawa, M., D. A. Nusinow, et al. (2007). "FKF1 and GIGANTEA complex formation is required for day-length measurement in Arabidopsis." Science **318**(5848): 261-265.
Schellmann, S., A. Schnittger, et al. (2002). "TRIPTYCHON and CAPRICE mediate lateral inhibition during trichome and root hair patterning in Arabidopsis." The EMBO journal **21**(19): 5036-5046.
Schittenhelm, S., U. Menge-Hartmann, et al. (2004). "Photosynthesis, carbohydrate metabolism, and yield of phytochrome-B-overexpressing potatoes under different light regimes." Crop Science **44**(1): 131-143.
Schnable, P. S., D. Ware, et al. (2009). "The B73 maize genome: complexity, diversity, and dynamics." Science **326**(5956): 1112-1115.
Seo, H. S., J. Y. Yang, et al. (2003). "LAF1 ubiquitination by COP1 controls photomorphogenesis and is stimulated by SPA1." Nature **423**(6943): 995-999.
Seo, M., A. Hanada, et al. (2006). "Regulation of hormone metabolism in Arabidopsis seeds: phytochrome regulation of abscisic acid metabolism and abscisic acid regulation of gibberellin metabolism." The Plant journal : for cell and molecular biology **48**(3): 354-366.
Sessa, G., M. Carabelli, et al. (2005). "A dynamic balance between gene activation and repression regulates the shade avoidance response in Arabidopsis." Genes Dev **19**(23): 2811-2815.
Sharrock, R. A. and T. Clack (2002). "Patterns of expression and normalized levels of the five Arabidopsis phytochromes." Plant Physiology **130**(1): 442-456.

Sharrock, R. A. and P. H. Quail (1989). "Novel phytochrome sequences in Arabidopsis thaliana: structure, evolution, and differential expression of a plant regulatory photoreceptor family." Genes Dev **3**(11): 1745-1757.
Sheehan, M. J., P. R. Farmer, et al. (2004). "Structure and expression of maize phytochrome family homeologs." Genetics **167**(3): 1395-1405.
Shen, H., L. Zhu, et al. (2008). "Light-induced phosphorylation and degradation of the negative regulator PHYTOCHROME-INTERACTING FACTOR1 from Arabidopsis depend upon its direct physical interactions with photoactivated phytochromes." Plant Cell **20**(6): 1586-1602.
Shen, Y., R. Khanna, et al. (2007). "Phytochrome induces rapid PIF5 phosphorylation and degradation in response to red-light activation." Plant Physiol **145**(3): 1043-1051.
Shimizu, T., A. Toumoto, et al. (1997). "Crystal structure of PHO4 bHLH domain-DNA complex: flanking base recognition." The EMBO journal **16**(15): 4689-4697.
Shin, J., K. Kim, et al. (2009). "Phytochromes promote seedling light responses by inhibiting four negatively-acting phytochrome-interacting factors." Proc Natl Acad Sci U S A **106**(18): 7660-7665.
Shin, J., E. Park, et al. (2007). "PIF3 regulates anthocyanin biosynthesis in an HY5-dependent manner with both factors directly binding anthocyanin biosynthetic gene promoters in Arabidopsis." Plant J **49**(6): 981-994.
Singh, A., M. T. Selvi, et al. (1999). "Sunlight-induced anthocyanin pigmentation in maize vegetative tissues." Journal of Experimental Botany **50**(339): 1619-1625.
Smith, H. (1995). "Physiological and Ecological Function within the Phytochrome Family." Annual Review of Plant Physiology and Plant Molecular Biology **46**: 289-315.
Smith, H. and G. C. Whitelam (1997). "The shade avoidance syndrome: Multiple responses mediated by multiple phytochromes." Plant Cell and Environment **20**(6): 840-844.
Smolen, G. A., L. Pawlowski, et al. (2002). "Dominant alleles of the basic helix-loop-helix transcription factor ATR2 activate stress-responsive genes in Arabidopsis." Genetics **161**(3): 1235-1246.
Somers, D. E., R. A. Sharrock, et al. (1991). "The hy3 Long Hypocotyl Mutant of Arabidopsis Is Deficient in Phytochrome B." Plant Cell **3**(12): 1263-1274.
Stapleton, A. E. and V. Walbot (1994). "Flavonoids can protect maize DNA from the induction of ultraviolet radiation damage." Plant physiology **105**(3): 881-889.
Sweere, U., K. Eichenberg, et al. (2001). "Interaction of the response regulator ARR4 with phytochrome B in modulating red light signaling." Science **294**(5544): 1108-1111.
Takano, M., N. Inagaki, et al. (2009). "Phytochromes are the sole photoreceptors for perceiving red/far-red light in rice." Proc Natl Acad Sci U S A **106**(34): 14705-14710.
Takano, M., N. Inagaki, et al. (2005). "Distinct and cooperative functions of phytochromes A, B, and C in the control of deetiolation and flowering in rice." Plant Cell **17**(12): 3311-3325.
Takano, M., H. Kanegae, et al. (2001). "Isolation and characterization of rice phytochrome A mutants." Plant Cell **13**(3): 521-534.
Tepperman, J. M., T. Zhu, et al. (2001). "Multiple transcription-factor genes are early targets of phytochrome A signaling." Proceedings of the National Academy of Sciences of the United States of America **98**(16): 9437-9442.
Tepperman, J. M., T. Zhu, et al. (2001). "Multiple transcription-factor genes are early targets of phytochrome A signaling." Proc Natl Acad Sci U S A **98**(16): 9437-9442.
Thiele, A., M. Herold, et al. (1999). "Heterologous expression of Arabidopsis phytochrome B in transgenic potato influences photosynthetic performance and tuber development." Plant Physiol **120**(1): 73-82.

Todaka, D., K. Nakashima, et al. (2012). "Rice phytochrome-interacting factor-like protein OsPIL1 functions as a key regulator of internode elongation and induces a morphological response to drought stress." Proc Natl Acad Sci U S A **109**(39): 15947-15952.
Toledo-Ortiz, G., E. Huq, et al. (2003). "The Arabidopsis basic/helix-loop-helix transcription factor family." The Plant cell **15**(8): 1749-1770.
Toledo-Ortiz, G., Y. Kiryu, et al. (2010). "Subcellular sites of the signal transduction and degradation of phytochrome A." Plant Cell Physiol **51**(10): 1648-1660.
Troyer, A. F. (2006). "Adaptedness and heterosis in corn and mule hybrids." Crop Science **46**(2): 528-543.
Ulijasz, A. T., G. Cornilescu, et al. (2010). "Structural basis for the photoconversion of a phytochrome to the activated Pfr form." Nature **463**(7278): 250-254.
Valdivia, E. R., D. Chevalier, et al. (2012). "DVL genes play a role in the coordination of socket cell recruitment and differentiation." Journal of experimental botany **63**(3): 1405-1412.
Vandenbussche, F., W. H. Vriezen, et al. (2003). "Ethylene and auxin control the Arabidopsis response to decreased light intensity." Plant Physiol **133**(2): 517-527.
Vandoren, M., A. M. Bailey, et al. (1994). "Negative Regulation of Proneural Gene Activity - Hairy Is a Direct Transcriptional Repressor of Achaete." Genes & Development **8**(22): 2729-2742.
Wada, T., T. Kurata, et al. (2002). "Role of a positive regulator of root hair development, CAPRICE, in Arabidopsis root epidermal cell differentiation." Development **129**(23): 5409-5419.
Wang, H. and X. W. Deng (2003). "Dissecting the phytochrome A-dependent signaling network in higher plants." Trends in plant science **8**(4): 172-178.
Wanner, L. A., G. Li, et al. (1995). "The phenylalanine ammonia-lyase gene family in Arabidopsis thaliana." Plant molecular biology **27**(2): 327-338.
Weatherwax, S. C., M. S. Ong, et al. (1996). "The interaction of light and abscisic acid in the regulation of plant gene expression." Plant physiology **111**(2): 363-370.
Wen, J., K. A. Lease, et al. (2004). "DVL, a novel class of small polypeptides: overexpression alters Arabidopsis development." The Plant journal : for cell and molecular biology **37**(5): 668-677.
Werker, E. (2000). "Trichome diversity and development." Advances in Botanical Research Incorporating Advances in Plant Pathology, Vol 31 2000 **31**: 1-35.
Whitelam, G. C. and P. F. Devlin (1997). "Roles of different phytochromes in Arabidopsis photomorphogenesis." Plant Cell and Environment **20**(6): 752-758.
Yamaguchi, R., M. Nakamura, et al. (1999). "Light-dependent translocation of a phytochrome B-GFP fusion protein to the nucleus in transgenic Arabidopsis." J Cell Biol **145**(3): 437-445.
Yang, J., R. Lin, et al. (2005). "Light regulates COP1-mediated degradation of HFR1, a transcription factor essential for light signaling in Arabidopsis." The Plant cell **17**(3): 804-821.
Zeier, J., B. Pink, et al. (2004). "Light conditions influence specific defence responses in incompatible plant-pathogen interactions: uncoupling systemic resistance from salicylic acid and PR-1 accumulation." Planta **219**(4): 673-683.
Zettl, R., J. Schell, et al. (1994). "Photoaffinity labeling of Arabidopsis thaliana plasma membrane vesicles by 5-azido-[7-3H]indole-3-acetic acid: identification of a glutathione S-transferase." Proceedings of the National Academy of Sciences of the United States of America **91**(2): 689-693.
Zhang, T., S. A. Maruhnich, et al. (2011). "Green light induces shade avoidance symptoms." Plant Physiol **157**(3): 1528-1536.
Zhang, W., Y. Sun, et al. (2006). "Regulation of Arabidopsis tapetum development and function by DYSFUNCTIONAL TAPETUM1 (DYT1) encoding a putative bHLH transcription factor." Development **133**(16): 3085-3095.
Zhu, Y., J. M. Tepperman, et al. (2000). "Phytochrome B binds with greater apparent affinity than phytochrome A to the basic helix-loop-helix factor PIF3 in a reaction requiring the PAS domain of PIF3." Proc Natl Acad Sci U S A **97**(24): 13419-13424.